

INN: SOBEK et al.

P. Date: 3/9/01

CRFE
SEARCH REQUEST FORM

U.S. DEPARTMENT OF COMMERCE
Patent and Trademark Office

Requestor's
Name: JANE ZARA

Serial
Number: 09/803,165

Date: 2-6-04

Phone: 2-0765

Art Unit: 1635

2718

11/81

Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors, keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

Please Search Seg ID No: 34

No size limits

Thanks

STAFF USE ONLY

Date completed: 2/10

Searcher: _____

Search Site

Vendors

STIC

IG

CM-1

STN

Pre-S

Dialog

Elapsed time: _____

Type of Search

APS

CPU time: _____

N.A. Sequence

Geninfo

Total time: _____

A.A. Sequence

SDC

Number of Searches: _____

Structure

DARC/Questel

Number of Databases: _____

Bibliographic

Other

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model1

Run on: February 9, 2004, 11:48:25 ; Search time 28 Seconds
(without alignment)
1169.593 Million cell updates/sec

Title: US-09-803-165-34

Perfect score: 4026

Sequence: 1 MIFDTDXITDGKPIIRFK.....KEDLKYQSSKQVGLDAWLRK 774

Scoring table: BLOSUM62

Gappen: 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3225.5	80.1	774	3 US-08-902-632-2	Sequence 2, Appli
2	3222.5	80.0	774	3 US-09-073-354-1	Sequence 1, Appli
3	3222.5	80.0	774	3 US-08-656-005A-1	Sequence 1, Appli
4	3222.5	80.0	774	3 US-09-073-259-1	Sequence 1, Appli
5	3222.5	80.0	774	3 US-09-363-195-1	Sequence 1, Appli
6	3222.5	80.0	774	3 US-09-418-027-1	Sequence 1, Appli
7	3209.5	79.7	778	2 US-08-906-925-4	Sequence 4, Appli
8	3151.5	78.3	775	1 US-07-966-778-1	Sequence 1, Appli
9	3151	78.3	775	1 US-08-424-921-1	Sequence 1, Appli
10	3151	78.3	775	2 US-08-556-355A-1	Sequence 1, Appli
11	3151	78.3	775	2 US-07-803-27A-1	Sequence 1, Appli
12	3151	78.3	775	2 US-09-244-889A-1	Sequence 1, Appli
13	3151	78.3	776	2 US-08-688-649-17	Sequence 1, Appli
14	3141.5	78.0	773	4 US-09-585-858-35	Sequence 35, Appli
15	3139	78.0	779	5 PCT-US95-152263-12	Sequence 12, Appli
16	3139	78.0	779	4 US-09-585-958-12	Sequence 32, Appli
17	2974	73.9	727	1 US-08-271-164A-8	Sequence 8, Appli
18	2502.5	62.2	1022	2 US-08-222-115B-27	Sequence 27, Appli
19	2502.5	62.2	1022	2 US-08-271-364A-7	Sequence 7, Appli
20	2069	51.4	1019	2 US-08-222-715B-26	Sequence 26, Appli
21	2069	51.4	1019	2 US-08-229-284A-2	Sequence 2, Appli
22	1595	39.6	396	1 US-08-907-166-6	Sequence 6, Appli
23	1403	34.8	788	4 US-09-391-340-6	Sequence 6, Appli
24	1403	34.8	788	2 US-08-907-166-4	Sequence 4, Appli
25	1194	29.7	803	4 US-09-391-340-4	Sequence 4, Appli
26	1194	29.7	803	1 US-08-062-368-4	Sequence 4, Appli
27	1191	29.6	803	1 US-08-062-368-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-08-902-632-2
; Sequence 2, Application US/08902632
; Patent No. 6008025
; GENERAL INFORMATION:
; APPLICANT: KOMATSUBARA, Hideyuki
; APPLICANT: KITABAYASHI, Masao
; APPLICANT: KANTIMURA, Hideki
; APPLICANT: KAWAKAMI, Bunsei
; APPLICANT: KAWAMURA, Yoshihisa
; APPLICANT: TAKAGI, Masahiro
; APPLICANT: IMANAKA, Tadayuki
; TITLE OF INVENTION: Modified Thermostable DNA Polymerase, and DNA Polymerase Composition for Nucleic Acid
; TITLE OF INVENTION: and DNA Polymerase Composition for Nucleic Acid
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenyon & Kenyon
; STREET: 1 Broadway
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.50 inch floppy disk, 1.44 Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 6.1 Windows
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/902/32
; FILING DATE: Concurrent Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 198911/96
; FILING DATE: 29-JUL-1996
; APPLICATION NUMBER: JP 200446/96
; FILING DATE: Concurrent Herewith
; APPLICATION NUMBER: 30-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: GREASON, Edward W.
; REGISTRATION NUMBER: 18,918
; REFERENCE/DOCKET NUMBER: 2418/7
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-425-7200
; TELEFAX: 212-425-5288
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 774 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein

418 VSPDTLNREGCKEYDVAPQVGHRFCKDFPGTIPSLLGDLLEERQKIKKGMKATIDPIERK 477
 481 MLDYRQAVKLUHNSYYGTYGPKARWYSKCEAESVTANGRYHLEMTEIEBKFGFKYL 540
 478 LDYRQAVKILANSTYYGPKARWYCEAESVTAGREYIIMTIEIEBKFGFKYL 537
 541 YADTDGFTATPGKETIICKAKEPKIYINSKJPGLLELEYEGFYLRGPFVAKRKYAVI 600
 538 YADTDGFTATPGKADATVCKMPEUNYNAKLPGALELEYEGFYKRGPFVTRKRYAVI 597
 601 DEEGRITTRGLEYRDWSEETAKETOAKVLFALLKEDSVKAVEIYKDVTEELAKYQFL 660
 598 DEEGRITTRGLEYRDWSEETAKETOAKVLFALLKEDSVKAVEIYKDVTEELAKYQFL 657
 661 EKLVHQITQDLSYEKAIGPHVAKRLANGKIKYRPGTISYVLRGGKISDRVIL 720
 658 EKLVHQITQDLSYEKAIGPHVAKRLANGKIKYRPGTISYVLRGGKISDRVIL 717
 721 SEMDPKHKYDDYIENQVPAVLRLAEGYRKEKDLYQSSKQVGLDAWLK 773
 718 DEEDEPTKHKYDAFYIENQVLPAAVERLRAFGYRKEKDLYQKTRQVGLSAWLK 770
 RESULT 3
 US 08-656-005A-1
 / Sequence 1, Application US/08656005A
 / Patent No. 6054301
 / GENERAL INFORMATION:
 / APPLICANT: KITABAYASHI, Masao
 / APPLICANT: ARAKAWA, Taku
 / APPLICANT: INOUE, Hiroaki
 / APPLICANT: KAWAKAMI, Bunsei
 / APPLICANT: KAWANTURA, Yoshihisa
 / APPLICANT: IMANAKA, Tadayuki
 / APPLICANT: TAKAGI, Masahiro
 / APPLICANT: Morigawa, Masakazu
 / TITLE OF INVENTION: A Method of Amplifying Nucleic
 / NUMBER OF SEQUENCES: 16
 / NUMBER OF SEQUENCES: 16
 / ADDRESS/AGENT ADDRESS:
 / ADDRESS: Kenyon & Kenyon
 / STREET: 1025 Connecticut Avenue, N.W., Suite 600
 / CITY: Washington
 / STATE: D.C.
 / COUNTRY: USA
 / ZIP: 20036
 / COMPUTER READABLE FORM:
 / MEDIUM TYPE: 3.5" Floppy disk
 / COMPUTER: IBM PC compatible
 / OPERATING SYSTEM: PC-DOS/MS-DOS
 / SOFTWARE: WordPerfect 5.1
 / CURRENT APPLICATION DATA:
 / APPLICATION NUMBER: US/08/656,005A
 / FILING DATE: 24 MAY 1996
 / PRIOR APPLICATION DATA:
 / APPLICATION NUMBER: JP 134096/95
 / FILING DATE: 31 MAY 1995
 / ATTORNEY/AGENT INFORMATION:
 / NAME: Toffenetti, Judith L.
 / REGISTRATION NUMBER: 39,048
 / REFERENCE/DOCKET NUMBER: 2418/3
 / TELECOMMUNICATION INFORMATION:
 / TELEPHONE: 202-429-1776
 / TELEX/FAX: 202-429-0796
 / INFORMATION FOR SEQ ID NO: 1:
 / SEQUENCE CHARACTERISTICS:
 / LENGTH: 774 amino acids
 / TYPE: amino acid
 / STRANDEDNESS: double
 / TOPOLOGY: linear
 / MOLECULE TYPE: protein
 / US-08-656-005A-1

Query Match 80.0%; Score 3222.5; DB 3; Length 774;
 Best Local Similarity 77.4%; Pred. No. 2e-242; Index 3; Gaps 2;
 Matches 598; Conservative 93; Mismatches 89; Indels 3;
 1 MTDITDITKDGKPIIRIFCKENGFXEYDHPFQPYIYALKDDSAIDEIKAIGSERHG 60
 1 MLDITDITEDGPVIRIFCKENGFXEYDHPFQPYIYALKDDSAIDEIKAIGSERHG 60
 61 KIVRVDVAKVCKKFLGRDVEWKLIFEFHPDVPVPAKGKTRHPAVIDYDIPPAKRY 120
 61 TVTVRVRVKEVQKFLGRPVEWKLIFTHPQVPARDKTRHGAVIDYDIPPAKRY 120
 121 LIDKG1PMEGDDEBLKUMAFDIEETYHGDEFGKGETIMSYADEBEARVITWKNDLDPY 180
 121 LIDKG1PMEGDDEBLKUMAFDQTLVHGGEEAEGPLMISADEBEARVITWKNDLDPY 180
 181 VDVYNSNEREMIKRFVQIVREKOPDVLITYNGDNFDLDPYIKAEGKLVGUTLIGRDEKEHPE 240
 181 VDVYSTEREMIKRFVYKEKDOPDVLITYNGDNFDAYLKERCEKGINGNFGDGS-E 238
 241 PKIHRMGDSFAVBIKGRIHFDFPVTIRINPFTYLEAVAVLGGTKSKGAAEIAAI 300
 239 PKIQRMGDRFAYEVKGRHFIDLYPVIRTINTNPFTYLEAVFQPKEXKYAAEITPA 298
 301 WETBESMRKLAQYSMEDRATYLGKFFPMBAELAKLIGQSVMDVSRSSTGNLVENYLL 360
 299 WEIGENLBRVARYSMEADAKVYELGKFFLPMQASLIGQSLWVDSRSSTGNLVENYLL 358
 361 RVAYERNELAPNKPDEEFYRRLRTYLGKFFPERGLWENITYLDFRCLYPSITVTHN 420
 359 RKAYERNELAPNKPDEKLARR PQSEGGTYKEPERGLWENIVYDFRSLYPSITVTHN 417
 421 VSPDTLREGCKNQVDAPIVGKFCDFGFFIPSISLGELITMREQEKKMKATIDPIEK 480
 418 VSPPLNREGCKEDVAQVGHFCDFGFFIPSISLGELITMREQEKKMKATIDPIEK 477
 478 LLDRQRATKILANSYYGGYGTARARWYCKCEAESVTAWGRHYTITKEIBKYGFKV 537
 541 YADTDGFTATPGKETIICKAKEPKIYINSKJPGLLELEYEGFYLRGFFYAKKYYAVI 600
 538 YSDIDGFFATPGDADETYVKKGMFELNYINAKLPGALELEYEGFYLRGFFYTKKCYAVI 597
 601 DEEGRITTRGLEYVRDWSIAKETQAKVLEALLEDPSVKAETVQDVBELAKYQVPL 660
 598 DEEGRITTRGLEYVRDWSIAKETQARVLEAKLKDSDVEAKRIVTEYERLKSKEVPP 657
 661 EKLYTHEQITKDLSEYKAIGPHVAKRLAAGKIKYRPGTISYVLRGSKGKISDRVILL 720
 658 EKLYTHEQITRDLKDYKATGPHVAVKLAARGVKIRPGTVISYVLRKGSGRIGDAIPF 717
 RESULT 4
 US-09-07-259-1
 / Sequence 1, Application US/09073259
 / Patent No. 6143536
 / GENERAL INFORMATION:
 / APPLICANT: IMANAKA, Tadayuki
 / APPLICANT: TAKAGI, Masahiro
 / APPLICANT: MORIKAWA, Masaki
 / TITLE OF INVENTION: DNA Encoding A Thermostable DNA Polymerase
 / NUMBER OF SEQUENCES: 16
 / CORRESPONDENCE ADDRESS:
 / ADDRESSEE: Kenyon & Kenyon
 / STREET: 1025 Connecticut Avenue, N.W., Suite 600
 / CITY: Washington
 / STATE: D.C.

COUNTRY: USA
ZIP: 20036
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 6.1 Windows
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/073,259
FILING DATE: Concurrent Herewith
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/656,005
FILING DATE: 24 MAY 1996
APPLICATION NUMBER: JP 134096/95
FILING DATE: 31 MAY 1995
ATTORNEY/AGENT INFORMATION:
NAME: Toffenetti, Judith L.
REGISTRATION NUMBER: 39,048
REFERENCE/DOCKET NUMBER: 2418/10
TELEPHONE: 202-429-1776
TELEFAX: 202-429-0796
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 774 amino acids
TYPE: amino acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-073-259-1

Query Match 80.0%; Score 3222.5; DB 3; Length 774;
Best Local Similarity 77.4%; Pred. No. 2e-242;
Matches 598; Conservative 83; Mismatches 89; Indels 3; Gaps 2;

Qy 1 MIFDTXYITRDGKPIIRIFKENGKFEKIDTDPHFOPIYIALLKDDSAIDBIKAIGERHG 60
1 MILDITYITEDGKPIIRIFKENGKFEKIDTDPHFOPIYIALLKDDSAIBEVKITAERHG 60
Qy 61 KIVRUVDAVKVKRKEKFLGRDVEWKLIIFHPODPVPAARGKIREHPIADYEDIPFKRY 120
61 TVTVTRVERVKQVKFLGRDVEWKLYTHPQDVPARDKIREHGAVIDYEDIPFKRY 120
Db 121 LIDKGLPIKPEGDEELKLMADFIDTYPHEGDBFGKGETIMISYADEBEEARVITWNIDLY 180
121 LIDKGLVMEGDEELKLMADFIDTYPHEGDBFGKGETIMISYADEBEEARVITWNIDLY 180
Qy 181 VDVSNSNREMIMKPFVQIVBEPDKDPLVITYGNDNFDLFLYIKAEKLGVTLLGRDKENHE 240
181 VDVSTTERMIMKPFVQIVBEPDKDPLVITYGNDNFDLFLYIKAEKLGVTLLGRDKENHE 240
Db 241 PKIHMGSDFAVBKGRHIDLEPVTARTINPFTYIPEAVENLQKTKSLGAEIAAI 300
239 PKIQMGSDFAVBKGRHIDLEPVTARTINPFTYIPEAVENLQKTKSLGAEIAAI 300
Qy 301 WETBSSMKLACKSMDARATYFLGKEFPMBAELAKLIGOSWVDSRSSTGNLVEYLL 360
301 WETBSSMKLACKSMDARATYFLGKEFPMBAELAKLIGOSWVDSRSSTGNLVEYLL 360
Db 361 RVAYERNELAPNKEDEEYRRRLRTYLLGGYKEPEGLWNTYLDFLRCLYPSLIVTHN 420
359 RKAYERNELAPNKEDEKELARR-RQSTEGGYKEPERGLWENIVDFRSIYPSLITEN 417
Qy 421 VSPDILREBECKNYDAPLVGYZKFCDPGFGFIPSILGLBLTRQEIKKRATIDPIKK 480
418 VSPDILREBECKNYDAPLVGYZKFCDPGFGFIPSILGLBLTRQEIKKRATIDPIKK 477
Db 481 MLDRQRAVKLHNSYTGNGYKARWYSKCERASVTAWGRTYIEMTIKEBEKFGFKV 540
478 LLDRQRAKILANSYTGNGYKARWYCKEAEASVTAWGRTYIEMTIKEBEKFGFKV 537
Qy 538 YSDTGFFFATPGADAEYTRKAMFPLNLYALPGLAEBLEYEGFVTRKICAVI 597
Qy 601 DEEGGTTTGRLEIYRDWSEIAKETQAKYLEAILEDSEVKAEEIVDVERBIAKYQVPL 660
Db 598 DEEGGTTTGRLEIYRDWSEIAKETQARVLEAILEDSEVKAEEIVDVERBIAKYQVPL 657
Qy 661 EKLVTHEQTRDKDISEYKAGPHVAAKLAAGKIKYRGTISIYVILRGSKISIDRVILL 720
Db 658 EKLVTHEQTRDKDISEYKAGPHVAAKLAAGKIKYRGTISIYVILRGSKISIDRVILL 717
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 718 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Db 719 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770

RESULT 5
US-09-363-095-1
Sequence 1, Application US/09363095
; Patent No. 6187513
; GENERAL INFORMATION:
; APPLICANT: IMANAKA, Tadayuki
; APPLICANT: TAKAGI, Masahiro
; APPLICANT: MORIKAWA, Masaki
; TITLE OF INVENTION: DNA Encoding A Thermostable DNA Polymerase
; FILE REFERENCE: 2418/11
; CURRENT APPLICATION NUMBER: US/09/363,095
; CURRENT FILING DATE: 1999-07-30
; EARLIER APPLICATION NUMBER: US 09/073,259
; EARLIER FILING DATE: 1998-05-06
; EARLIER APPLICATION NUMBER: US 08/656,005
; EARLIER FILING DATE: 1996-05-24
; EARLIER APPLICATION NUMBER: JP 134096/95
; EARLIER FILING DATE: 1995-05-31
; NUMBER OF SEQ ID NOS: 16
; SEQ ID NO: 1
; SOFTWARE: WordPerfect 6.1 Windows
; LENGTH: 774
; TYPE: PRT
; ORGANISM: Hyperthermophilic archaeon
; US-09-363-095-1

Query Match 80.0%; Score 3222.5; DB 3; Length 774;
Best Local Similarity 77.4%; Pred. No. 2e-242;
Matches 598; Conservative 83; Mismatches 89; Indels 3; Gaps 2;

Qy 1 MIFDTXYITRDGKPIIRIFKENGKFEKIDTDPHFOPIYIALLKDDSAIDBIKAIGERHG 60
1 MILDITYITEDGKPIIRIFKENGKFEKIDTDPHFOPIYIALLKDDSAIBEVKITAERHG 60
Qy 61 KIVRUVDAVKVKRKEKFLGRDVEWKLIIFHPODPVPAARGKIREHPIADYEDIPFKRY 120
121 LIDKGLVMEGDEELKLMADFIDTYPHEGDBFGKGETIMISYADEBEEARVITWNIDLY 180
Db 121 LIDKGLVMEGDEELKLMADFIDTYPHEGDBFGKGETIMISYADEBEEARVITWNIDLY 180
Qy 181 VDVSNSNREMIMKPFVQIVBEPDKDPLVITYGNDNFDLFLYIKAEKLGVTLLGRDKENHE 240
181 VDVSTTERMIMKPFVQIVBEPDKDPLVITYGNDNFDLFLYIKAEKLGVTLLGRDKENHE 240
Db 241 PKIHMGSDFAVBKGRHIDLEPVTARTINPFTYIPEAVENLQKTKSLGAEIAAI 300
239 PKIQMGSDFAVBKGRHIDLEPVTARTINPFTYIPEAVENLQKTKSLGAEIAAI 300
Qy 301 WETBSSMKLACKSMDARATYFLGKEFPMBAELAKLIGOSWVDSRSSTGNLVEYLL 360
301 WETBSSMKLACKSMDARATYFLGKEFPMBAELAKLIGOSWVDSRSSTGNLVEYLL 360
Db 361 RVAYERNELAPNKEDEEYRRRLRTYLLGGYKEPEGLWNTYLDFLRCLYPSLIVTHN 420
359 RKAYERNELAPNKEDEKELARR-RQSTEGGYKEPERGLWENIVDFRSIYPSLITEN 417
Qy 421 VSPDILREBECKNYDAPLVGYZKFCDPGFGFIPSILGLBLTRQEIKKRATIDPIKK 480
418 VSPDILREBECKNYDAPLVGYZKFCDPGFGFIPSILGLBLTRQEIKKRATIDPIKK 477
Db 481 MLDRQRAVKLHNSYTGNGYKARWYSKCERASVTAWGRTYIEMTIKEBEKFGFKV 540
478 LLDRQRAKILANSYTGNGYKARWYCKEAEASVTAWGRTYIEMTIKEBEKFGFKV 537
Qy 538 YSDTGFFFATPGADAEYTRKAMFPLNLYALPGLAEBLEYEGFVTRKICAVI 597
Db 598 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 601 DEEGGTTTGRLEIYRDWSEIAKETQAKYLEAILEDSEVKAEEIVDVERBIAKYQVPL 660
Db 658 EKLVTHEQTRDKDISEYKAGPHVAAKLAAGKIKYRGTISIYVILRGSKISIDRVILL 720
Qy 661 EKLVTHEQTRDKDISEYKAGPHVAAKLAAGKIKYRGTISIYVILRGSKISIDRVILL 717
Db 718 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 719 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 718 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 719 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 718 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 719 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 718 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 719 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 718 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 719 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 718 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 719 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 718 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 719 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 718 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 719 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 718 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 719 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 718 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 719 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 718 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 719 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 718 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 719 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 718 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 719 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 718 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 719 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 718 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 719 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 718 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 719 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 718 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 719 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 718 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 719 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 718 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 719 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 718 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 719 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 718 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 719 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 718 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 719 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 718 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 719 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 718 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 719 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 718 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 719 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 718 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 719 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 718 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 719 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 718 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 719 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 718 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 719 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 718 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 719 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 718 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 719 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 718 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 719 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 718 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 719 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 718 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 719 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 718 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 719 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 718 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 719 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 718 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 719 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 718 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 719 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 718 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 719 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 718 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 719 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 718 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 719 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 718 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 719 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 718 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 719 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 718 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 719 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 718 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 719 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 718 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 719 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773
Db 718 DEFDPTKHKYDAEYTYIENQVLPAVIRLAEFGYKEDLRYQTRQVGLISANLK 770
Qy 721 SEYDPKKHCKYDPDXYIENQVLPAVIRLAEFGYKEDLKYQSSKQYGLDANK 773

RESULT 6
 US-09-418-027-1
 Sequence 1, Application US/09418027
 Patent No. 6225055
 GENERAL INFORMATION:
 APPLICANT: KITABAYASHI, Masao
 APPLICANT: ARAKAWA, Taku
 APPLICANT: INOU, Hiroaki
 APPLICANT: KAWAKAMI, Bunsei
 APPLICANT: IMANAKA, Tadayuki
 APPLICANT: TAKAGI, Masahiro
 APPLICANT: MORIKAWA, Masaaki
 TITLE OF INVENTION: A Thermostable DNA Polymerase and Kits for
 TITLE OF INVENTION: Amplifying Nucleic Acids
 NUMBER OF SEQUENCES: 16
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Kenyon & Kenyon
 STREET: 1025 Connecticut Avenue, N.W., Suite 600
 CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20036
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC DOS/MS-DOS
 SOFTWARE: Wordperfect 6.1 Windows
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/418,027
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 09/073,354
 FILING DATE: 31 MAY 1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Toffenett, Judith L.
 REGISTRATION NUMBER: 39,048
 REFERENCE/DOCKET NUMBER: 134096/95
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-429-1776
 TELEX/FAX: 202-429-0796
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 774 amino acids

TYPE: amino acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: Protein
 US-09-418-027-1
 Query Match 80.0%; Score 3222.5; DB 3; Length 774;
 Best Local Similarity 77.4%; Pred. No. 2e-242;
 Matches 598; Conservative 83; Mismatches 89; Indels 3; Gaps 2;
 1 MIFDTUYITDKGPTRIFPKKENGEKIEFLDPHPFQTYIAALLKDSAIDIKAKGERIG 60
 1 MILDIDVITEDGKPVTRIFPKKENGEKIEFLDTPFQTYIAALLKDSAIEFVKTAERHG 60
 61 KIVRYVDIVKVKKKFLGRDVEWKLUJFHEHQDVPALRKTREHDAVIDYBYDIFPAKY 120
 61 TVVTGRVERVYQKKFLGRPVWVKQYFTHQDVPARDKIREHGVIVDYEYDIFPAKY 120
 121 LIDKGLIPMEGDEEKLMLAFDIEFTYHEGDEFGKGEIIMSYADEBEARVITWNIDLPY 180
 121 LIDKGLIPMEGDEEKLMLAFDQTYHEGEFAEPILMISYADEBEARVITWNIDLPY 180
 181 VDVVSNEREMTKRFTQVIVREKDPLVITNGDNFLPILKRAKNGTLLGRKREHPE 240
 181 VDVVSNEREMTKRFLRVVKEKDPLVITNGDNFLPILKGRKREKLGINFALGRGS--E 238
 241 PKIHRNGDSFAVEIIGRIRHDPFLPVVRRTINLPYTLAEVAGLKTKSGLGAEIAAI 300
 239 PKIQRNGDSFAVEVGRIRHDPFLPVVRRTINLPYTLAEVAGLKTKSGLGAEIAAI 298
 301 WETEESMKLQLQSYMEDARATYELGKEFFMEAEIYLQGQSVDVSRSSTGNLVEWYLL 360
 299 WETGENLERYARYSMEDAKTYELKEFPLMEAQSLRQLGSLNWVRSRSSTGNLVEWYLL 358
 361 RVAYERNELAQNPKDEEYERRLRTYLGQYXKEPERGWIWENTYLDFFRLYPSLIVTHN 420
 359 RKAYERNELAQNPKDEKLLR - RSYEGSYVKEPERGWIWENTYLDFFRLYPSLIVTHN 417
 421 VSPDTLREGCKKNQDVAPIVYKFCDFPGFIPSTLGEIITMRQEIJKKOMKATIDPIEK 480
 418 VSPDTLNRGCKKEYDVAPOVGHFRCDPFPFIPSGLDQEERQKCKMKATIDPIEK 477
 481 MLDYQRRAVKHLANSYYGMYGPKARWYSKECAESTVANGRHYIEMTIKEEKFQKVY 540
 478 LLDYQRRAKILANSYYGYYARWYCKECAESTVANGREY-TMTIKEEKFQKVY 537
 541 YADDDFYATIPEGREPETIKKAKFLK2INSKUGPLLLEYEGYLRGPFAKRYAVI 600
 538 YSDTDOFFATIPEGDAETVKKAKFENLYNIAKLPGALLEYESFYKRGFFFVTKKYAVI 597
 601 DEEGRITTRGDEVVDRWSIAKETOQAKYEAILEDYSVKAVIYDVEBIAKRYQYPL 660
 598 DEEGKTRITRGLTEVDRWSIAKETOQAKYEAILEDYSVKAVIYDVEBIAKRYQYPL 657
 661 EKLVTHEQITKMDLSYKAICGPHVATAKRLAAGKIKVRPGTIISYVLRSGKTSDRVILL 720
 658 EKLVTHEQITRDLKQYKATGPHVATAQKLLARGYKIRPTVISVLRGSGRIGDRAI PP 717
 721 SEYDPKHKYDQDYYIENQVLPAVIRLAFGKREKDIKYQSSSKQVGDAMWK 733
 718 DBFDPTKHKYDAEYTTENQVLPAVIRLAFGKREKDIKYQSSSKQVGDAMWK 770

RESULT 7
 US-08-906-925-4
 Sequence 4, Application US/08906925
 Patent No. 5982904
 GENERAL INFORMATION:
 APPLICANT: Redl, William
 APPLICANT: Fly, Susan
 APPLICANT: Kapoor, Barbara F.
 APPLICANT: Nye, Steven H.
 APPLICANT: Ting, Eve

TITLE OF INVENTION: THERMOCOCCUS BAROSII DNA POLYMERASE MUTANTS
 NUMBER OF SEQUENCES: 9
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Charles E. Brady
 STREET: 411 East Wisconsin Avenue
 CITY: Milwaukee
 STATE: Wisconsin
 COUNTRY: U.S.A.
 ZIP: 53202-4497
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/906,925
 FILING DATE:
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Baker, Jean C.
 REGISTRATION NUMBER: 35 433
 REFERENCE/DOCKET NUMBER: 740211.90628
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (414) 277-5703
 TELEFAX: (414) 271-3552
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 778 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: Protein
 US-08-906-925-4

Query	Match	Score	Length
Qy	1 MIEDDTIKDGGPKIIRPKKENGEKIELDPHPQQPYIALLKDDAIDEKAIK 1 MILDDVY-TEDGKPVIRPKDKGKFGKLYDREFEPYIALLRDDAIEELKIK	3209.5	778
Ds	596; Best Local Similarity 79.7%; Pred. No. 2.1e-24; Matches 83; Mis matches 90; Indexes 3; C		
Qy	61 KIVRVDAVKKKKFLGRDVEVKKLIFHPQDVPAIROLRKIREHPAVDIDYEDIP 61 KVVKVKRAEKVKKKFLGRSVEVWLYFTHPQDVPAIDKIRGHPAVDIDYEDIP		
Ds	121 LIDKGLIPMEGDBELKOMAFDIEFTYHEGDFERGEKIMISTYADEEEARVITWKN 121 LIDKGKLVMEGDBEELKLMFSDIEFTYHEGEEFOTGPILMISTADEESEARVITWKN		
Qy	181 VDVVSNEREMIKRFVQIVREKDPDVLLITYNGDNFDLPLVYKAKGTYLLGRD 181 VDVVSTEKMIKFLKVTKERDDVLLITYNGDNFDAYLKRCELGVSFTLGRD		
Ds	241 PKTHRGDSFAVBIKGRTHFDLPPVYRTEINLPTTYLEAVYFAVLGKTKSKGIAE 239 PKLQRMGDRFAVVKGR-HFDLPPVYRTEINLPTTYLEAVYFAVLGKTKSKGIAE		
Qy	301 WETEESMKKLAQYSMEDARATYELGKERRPPMEEAELAKLIGOSWYDVSRSSTGNLV 299 WETSEGLRVARYSMEDARAVTYELGREFPMEAQSLIGQGLWDVSRSSTGNLV		
Ds	361 RYAYERNELAPNPKDDEEYRRRTTLLGGTYKEPERLWENITYLDFCLYPSI 359 RYAYERNELAPNPKDDELLARR-RGGAGGYKEPERGSRWDNIVYDFRSLYPSI		
Qy	421 VSPDTLREGCNYDVAFLVYKFCDKPGFTPSLGLSLQSLIGQGLWDVSRSSTGNLV 418 VSPDTLNEGCKSYDVAQVGHFKCDPGFTSLGLNLEROKIKRKMATID		
Ds	481 MLDYRQRAVKLHANSYYGMGYPKARMYSKCAESVATAGHYIEMTIKEIbekP 478 TADYDOPKATTAATANSPEVYVCPDPAVCKPFTSSTWCPSEVFTMTPYKVEY		

RESULT 8

US-07-966-278-1

/ Sequence 1, Application US/07966278

/ Patent No. 5489523

/ GENERAL INFORMATION:

/ APPLICANT: Mathur, Eric A

/ TITLE OF INVENTION: EXONUCLEASE-DEFICIENT THERMOSTABLE

/ NUMBER OF SEQUENCES: 4

/ TITLE OF INVENTION: PYROCOCCUS FURIOSUS DNA POLYMERASE I

/ CORRESPONDENCE ADDRESS:

/ ADDRESS: Thomas Fitting

/ STREET: 12526 High Bluff Road, Suite 300

/ CITY: San Diego

/ STATE: CA

/ COUNTRY: USA

/ ZIP: 92130

/ COMPUTER READABLE FORM:

/ MEDIUM TYPE: Floppy disk

/ COMPUTER: IBM PC compatible

/ OPERATING SYSTEM: PC-DOS/MS-DOS

/ SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

/ APPLICANT NUMBER: US/07/966,278

/ FILING DATE: 1991-12-26

/ CLASSIFICATION: 435

PRIOR APPLICATION DATA:

/ APPLICATION NUMBER: 1

FILING DATE:

ATTORNEY/AGENT INFORMATION:

/ NAME: Fitting, Thomas S

/ REGISTRATION NUMBER: 34,163

/ REFERENCE/DOCKET NUMBER: STG0113P

TELECOMMUNICATION INFORMATION:

/ TELEPHONE: 619-792-3680

/ TELEFAX: 619-792-8477

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

/ LENGTH: 775 amino acids

/ TYPE: AMINO ACID

/ TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ANTI-SENSE: NO

US-07-966-278-1

Query Match 78 3%; Score 3151; DB 1; Length 775;

Best Local Similarity 74 6%; Pred. No. 7-4e-237;

Matches 576; Conservative 100; Mismatches 94; Indels 2; Gaps 1;

Qy 1 MIFDTDYITKDGPKIIRIFPKKENGGPKIELDPHFOPTYTALLQDSDAIDETIKAIKGGERHG 60

Db 1 MILDVDYITEGGPKVIRIFPKKENGGPKIELDRTFFPYTALLRDSKIEPVKKTGERHG 60

Qy 61 KIVRIVYDAVKYKKFLGRDYEWKLJFEHPODQVPLRGKIREHPAVIDYEDIFPAKRY 120

Qy 541 YADTDFGTYATPGEPEPTEKKKAKFLKTYINSKGLLELEYGGYLRLGFVAKRYAVI 600

Db 538 YADTDFGTYATPGEPEPTEKKKAKFLKTYINSKGLLELEYGGYLRLGFVAKRYAVI 597

Qy 601 DEEGRITTRGIVVWRDWEAKETQAKVLEAIIKEDSVEKAEVTKDVVEIAKXQVPL 660

Db 598 DEEGRITTRGIVVWRDWEAKETQAKVLEAIIKEDSVEKAEVTKDVVEIAKXQVPL 657

Qy 661 EKLVTHEQITKDLBSYKAIGPHVAKRLAAGKIKVPRGTTISYVLRGSGKISDRVILL 720

Db 658 EKLVTHEQITRELKDVKATGPHVAKRLAARGKIRGPTVISTVLKGSGRIGRAIPF 717

Qy 721 SEYDPKKHKYTDPEDXYTIENQVLPAVRILPAGYKEDLKYQSSKQGDAWL 772

Db 718 DEFPDKHRYTDAYTENQVLPAVRILPAGYKEDLRYQKTRGGLGAWL 769

61 KIVRIVDVEKVEKFLGKPTVWVWLYLEPQDVTIREKVRHAPAVDIFENDIPFAKRY 120
 121 LIDKGILIPMEDEELKLMADFETYHEGDEFKGKEIMISYADEEFARVITKWNIDLPY 180
 121 LIDKGILIPMEGEELKLMADFETYHEGEGFKGPIMISYADENEAKVITKWNIDLPY 180
 181 VDVVSBEREMKRFQIVREKDPDYLITNGNDPDLPLIKRAEKLGYTLLGRDEKHEPE 240
 181 VEVSSBERENKRFIIRKDPLIVTNGDSDFPLAKRAEKLGLKTIGRDGS--E 238
 241 PKIHRGDSPAVEIKGRIHEDLPLPVVRINLPPTYLAVYEAVLGKTKSKGAETAAI 300
 239 PRMQRGDMPAVERGRIRHDFLVHITRINLPPTYLAVYEAVLGKTKSKGAETAAI 298
 301 WETEESMMKLLAQYSEMDARATEYLGKEPPMEEALAKLJGQSYWVDSRSSTGNLNVWYLL 360
 299 WEGSEMLERVAYKSNMEDAKATYLGKEEFLPMEQLSRLVQDWPWDSRSSTGNLNVWENFLL 358
 361 RVAYEENELAPNPKDPEETIRRLLFTTYLGKEPERGLWENTYLDERCLYPSITVTHN 420
 359 RKAYEENEVAPNPKDPEETQRLFESYTGFGKEPERGLWENIVYLDFAIYPSITHN 418
 421 VSPDTLREGCKNDTDAPIVGYKFCDFGGFIPSIGLGLITMREIKKKKMKTIDPTEKK 480
 419 VSPDTLNLGCKNDTDAPIVGYKFCDFGGFIPSIGLGLITMREIKKKKMKTIDPTEKK 480
 481 MLDYRORAKLHANSYYGGMGYPKARYMSKECAESYTAWGRHYIEMTIKELEKFKFVKL 540
 479 LLDYRORAKLHANSYYGGMGYPKARYMSKECAESYTAWGRHYIEMTIKELEKFKFVKL 538
 541 YADTNGFYATIPGKPET1KKKAKPEFLKINSKLPGLELEFEGYLRGFVAKRKYAVI 600
 539 YDFTDGLYATIPGGESEET1KKKAKPEFLKINSKLPGLELEFEGYLRGFVAKRKYAVI 598
 601 DEEGRITLGLEVVRDNDSEIATQAKVLEA1LKEDSVKEAVIYRDVWEIAKQVPL 660
 599 DEEGRITLGLEVVRDNDSEIATQAKVLEA1LKEDSVKEAVIYRDVWEIAKQVPL 660
 661 EKLYIHEQITKDLBKA1GPHYAAKRLAAKCTKVRPGTIIISYVLRGSGKISDRVILL 720
 659 EKLAAYEQTPLRPHYKA1GPHYAAKRLAAKCTKVRPGMIVGYVLRGDPISRAILLA 718
 721 SEYDFKQKHYDPDYYIENQVPLAVRLLEAFGYRKEDLKYQSSKQGLDWWL 772
 719 EYDFKQKHYDAEYIENQVPLAVRLLEAFGYRKEDLRYORTQVGLTWSL 770

RESULT 9
 US - 08-424-921-1
 Sequence 1, Application US/08424921
 ; Patent No. 5545552
 ; GENERAL INFORMATION:
 ; APPLICANT: Mathur, Eric A
 ; TITLE OF INVENTION: PURIFIED THERMOSTABLE PYROCOCCUS
 ; TITLE OF INVENTION: FURIOUS DNA POLYMERASE I
 ; NUMBER OF SEQUENCES: 6
 ; CORRESPONDENCE ADDRESS: 6
 ; ADDRESSEE: Bingham & Fitting
 ; STREET: 12526 High Bluff Road, Suite 300
 ; CITY: San Diego
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 92130
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/424, 921
 ; FILING DATE: 19-APR-1995
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:

Db 59 DEEGKVITRGLIEVERDWEIAKETQARVDETLIHKGYDVEAIVKIVQKANYETLPP 658
 Qy 61 EKLVHEQITKDLSEBYKAIGPHVATAKRLAAKGKIVRPGTIISIVLFGSGKTSDRVILL 720
 Db 659 EKLAYEQTIRPLHYKAIGPHVATAKRLAAKGKIVRPGVIVLREDGPTSNRALIA 718
 Qy 721 SEYDPKKHKKYDQYYTENQVLPAVRLAEGYRKEDLYKQSSKQVGLDAWL 772
 Db 719 EYDPKCKHKKYDAEYYTENQVLPAVRLAEGYRKEDLYKQTRQVGLTSWL 770

RESULT 10
 US-08-556-355A-1
 / Sequence 1, Application US/08556355A
 / GENERAL INFORMATION:
 / APPLICANT: MATHUR, Eric J.
 / TITLE OF INVENTION: Thermostable Pyrococcus
 / TITLE OF INVENTION: furious DNA Polymerase I
 / NUMBER OF SEQUENCES: 6
 / CORRESPONDENCE ADDRESS:
 / ADDRESSEE: Finnean, Henderson, Farabow, Garrett &
 / STREET: Dunner, L.L.P.
 / CITY: Washington
 / STATE: D.C.
 / COUNTRY: USA
 / ZIP: 20005
 / COMPUTER READABLE FORM:
 / MEDIUM TYPE: Floppy disk
 / COMPUTER: IBM PC compatible
 / OPERATING SYSTEM: PC-DOS/MS-DOS
 / SOFTWARE: Patent in Release #1.0, Version #1.30
 / CURRENT APPLICATION DATA:
 / APPLICATION NUMBER: US/08/556,355A
 / FILING DATE: 13-NOV-1995
 / CLASSIFICATION: 435
 / PRIOR APPLICATION DATA:
 / APPLICATION NUMBER: US 08/424,921
 / FILING DATE: 19-APR-1995
 / PRIOR APPLICATION DATA:
 / APPLICATION NUMBER: US 07/803,627
 / FILING DATE: 02-DEC-1991
 / PRIOR APPLICATION DATA:
 / APPLICATION NUMBER: US 07/779,846
 / FILING DATE: 21-OCT-1991
 / PRIOR APPLICATION DATA:
 / APPLICATION NUMBER: US 07/657,073
 / FILING DATE: 19-FEB-1991
 / PRIOR APPLICATION DATA:
 / APPLICATION NUMBER: US 07/620,568
 / ATTORNEY/AGENT INFORMATION:
 / NAME: Barker, M. Paul
 / REGISTRATION NUMBER: 32,013
 / REFERENCE DOCKET NUMBER: 04121.0004-02
 / TELECOMMUNICATION INFORMATION:
 / TELEPHONE: (202) 408-4000
 / TELEFAX: (202) 408-4400
 / SEQUENCE CHARACTERISTICS:
 / LENGTH: 775 amino acids
 / TYPE: amino acid
 / STRANDEDNESS: single
 / TOPOLOGY: linear
 / MOLECULE TYPE: protein
 / US-08-556-355A-1

Db 1 MILDVYDITPEFGKPVRLFKENGFKIEDRTPPYTALLRDSKIEBVKITGERHG 60
 Qy 61 KIVRUYDIAVYKVKKKFGRDVWVKLIFEPDQVPLRGKTRHEDPVIDYBYDIFAKRY 120
 Db 61 KIVRUYDVEKVKFQKPTWVKYLEHQDVPTRKREHPPAVDIFEDIFARRY 120
 Qy 121 LIDKGLIPMGEDEELKLMADIEETYTHEGDEFGKCBIMISYADEDEARVITWKLMDLPY 180
 Db 121 LIDKGLIPMGEDEELKLMADIEETYHEGEGKOPIMISYADENEAKVITWKLMDLPY 180
 Qy 181 VDVSNSEREMKRTQIVREKDPDVLITNGDNFLPVLKRAFLGTVLGRKAEKHE 240
 Db 181 VEVVSSEREMKRFPLTIREKDPDVLITNGDSFPPFLAKRFLGKLTIGRGS-E 238
 Qy 241 PKIHRMGDSFAVEIKGRIFHDLPVPRTRINLPTLEAVEAVGKTKSKLGAEBIAAI 300
 Db 239 PKMQPFGDMTAVEVKGRIFHLYHTRTINLPTLEAVEATGKPKBKVYADEIAKA 298
 Qy 301 WETEESMKKLAQYSMEDARATEYLGKFFMBAELAKLIGQSVWVRSSTGNLVBWYLL 360
 Db 299 WESGENLERVAKYSMEDAKATYELGKFFMELQISRLVQPLWVRSSTGNLVBWYLL 358
 Qy 361 RYAYERNELAPKPDDEBEYRRLRTYLGGYVKPERGLMENITYLDFRLCLYPSLITIN 420
 Db 359 RYAYERNEVAPKPKSEPEYQRLRSYTGTVKEPEKGLMENIVVLDRLALYPSLITIN 418
 Qy 421 VSPDTLREGCKQNYDIAPIVGYKFCDFG3PTIPSIGELTMROEKKKOMATIDPIEK 480
 Db 419 VSPDTLNLLEGCKQNYDAPQVHKFKCD1GPTSLIGHJLUEERQKLTKMKETODIEKI 478
 Qy 481 MLDYRQBAVKLHANSYYGMYKPARWMSKCAESYTAWGRHYIEMTKEIIEKGFKVYL 540
 Db 479 LLDYRQAKIILANSPFGYCYAKWYCCEAESTAWGRKYIILWKEIIEKGFKVYL 538
 Qy 541 YADTDGFATIGERKPTIKKAKFELKYIINSKLQPLLELEYEGFLYLRGSPVAKRYAVI 600
 Db 539 YIDTDGLYATIGEGSEBEIKKAKFELKYIINSKLQPLLELEYEGFLYLRGSPVAKRYAVI 598
 Qy 601 DEBRGKTRGLEVRWOWSEIATEKTOAKVLEAILEDSEKAVEIYKDVYEEIAKYQVPL 660
 Db 599 DEEGKVTRGLEVRWOWSEIATEKQARVLETIILKHGDVEAARVLYKEIQLKAYEIPP 658
 Qy 661 EKLVTHEQTKDLSYEKAIGPHVATAKRLAAKGKIVRPGTISYVLRGSKSKISDEVILL 720
 Db 659 EKLAIYEQTIPHEYKAIGPHVAVAKLAAKGKIVRPGVIGYVLRGDGPISNRAILA 718
 Qy 721 SEYDPKKHKKYDQYYTENQVLPAVRLAEGYRKEDLYKQSSKQVGLDAWL 772
 Db 719 EYDPKCKHKKYDAEYYTENQVLPAVRLAEGYRKEDLYKQTRQVGLTSWL 770

RESULT 11
 / Sequence 1, Application US/07803627A
 / Patent No. 5948663
 / GENERAL INFORMATION:
 / APPLICANT: MATHUR, Eric J.
 / TITLE OF INVENTION: Purified Thermostable Pyrococcus
 / NUMBER OF SEQUENCES: 6
 / CORRESPONDENCE ADDRESS:
 / ADDRESSEE: Finnean, Henderson, Farabow, Garrett &
 / STREET: 1300 I Street, N.W.
 / CITY: Washington
 / STATE: D.C.
 / COUNTRY: USA
 / ZIP: 20005
 / COMPUTER READABLE FORM:
 / MEDIUM TYPE: Floppy disk
 / COMPUTER: IBM PC compatible
 / OPERATING SYSTEM: PC-DOS/MS-DOS

Query Match 78.3% Score 3151; DB 2; Length 775;
 Best Local Similarity 74.6% Pred. No. 7.4e-23;
 Matches 576; Conservative 100; Mismatches 94; Indels 2; Gaps 1;
 1 MIFDDYDITKDGKPLTRIFKENGFKIELDPHFPQPYIYALLKDDSAIDEIAKIKGERHG 60
 Qy

SOFTWARE: Patent In Release #1.0, Version #1.3.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/803,627A
 FILING DATE: 02-DEC-1991
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/776,552
 FILING DATE: 14-OCT-1991
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/657,073
 FILING DATE: 19-DEC-1991
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/620,568
 FILING DATE: 03-DEC-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: Barker, M. Paul
 REGISTRATION NUMBER: 32,013
 REFERENCES/DOCKET NUMBER: 04121.0004-000
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 408-4000
 TELEFAX: (202) 408-4400
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 775 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 US-07-803-627A:1

Query Match 77.3%; Score 3151; DB 2; Best Local Similarity 74.6%; Prcd. No. 7.4e-237; Matches 576; Conservative 100; Mismatches 94; 1

Qy 1 MIFDDYTKMDGKPLTRPKKENGEFKELDHPHQPYTYAII
 Db 1 MILDYDYEGRKVIRLPRKENKFKEHDRIFRPRTYIAY
 Qy 61 KIVRYVDAVKKKFLGRDVEVNLKLFLEHPHQDPAFLGKIR
 Db 61 KIVRIVDVERKPLGPITVVKLYLEHPHQDPTIREKVR
 Qy 121 LIDGKLIPMEGDEELKLMADFIEFYHAGDETRGKETIMISSY
 Db 121 LIDGKLIPMEGEEELKLMADFIELYHEGEBGKGPIMISY
 Qy 181 VDVSNEREMIKRQVQIVREKOPDVLITYNGDNFD-LYLYLKR
 Db 181 VEVHSERMRMFKRFLIREKOPDVLITYNGSFDPYLAKR
 Qy 241 PKIHRMGDSFAVEIKGRTHFDLFPVYVRTINLPTTYLEAVYEE
 Db 239 PRM0RIGDMTAVEKGRTHFDLTHVTRINTLPTTYLEAVYEE
 Qy 301 WETEESMKLQYSMEDARATYELGKRFEPMDNLKLLIGOS
 Db 359 RKAVERNEVAPNPSBEEYQRLLRESYTGPFYKEPKGLWNN
 Qy 361 RVAVERNELAPNPKDDEEMRRLRTYLGKRFEPMDNLKLLIGOS
 Db 421 VSPDTLREGCKNDVATIGKXFCRDFFGFFPSLIGELBLTM
 Qy 419 VSPDTLNLGCKNDVATIGKXFCRDFFGFFPSLIGLHLLE
 Db 481 MLDYRORAVKLHANSYYGKMPKARYWSKRCDAESTUAWGRH
 Qy 479 LLDYRQAIKLANSFGYYGAKARYWCKEGAESSTAWGRK
 Db 541 YADTGFYATIPIGKPKETIKKAKFELKYINSKLPOLELEY
 Db 539 YIDTGDGLATIPIGKESBIIKKALEYFQYINSKUPGLLEY

Matches 576; Conservative 100; Mismatches 94; Indels 2; Gaps 1;

Query 1 MIFDDYITDKGKPIRIFKKEPFKEFDLPHFQPYIALLYKDDSALEDEKAIAKGPRHG 60
 Database 1 MILDDYITCEGPKVIRPKKENKFKLEHDRTERPYYLRRDSKKEEVKTKTGERHG 60

Query 61 KIVRVDAAVKVKKPLGRDVEWKLIFLIFEPQDVPALRKIREPAVILYEDIPFKARY 120
 Database 61 KIVRVDVERKEVKPLGKPTVWKLLEHEDQVDTIREKREPAVILYEDIPFKARY 120

Query 121 LIDKGJLPPMGEDEKLMAFDIETTYHEDEDFGKGEIMISYADEEAVRVTWNIDPY 180
 Database 121 LIDKGJLPPMGEDEKLMAFDIETTYHEDEDFGKGEIMISYADEEAVRVTWNIDPY 180

Query 181 VDVVSNEREMIKRKYQIVREKDPDVLITYNGDNFDLPLVILKRAEKGTYLILLGRDKHPE 240
 Database 181 VEVVSSEREMIKRFLRIRKDPDVLITYNGDSDFPLAKRAEKGKLTIGRGS-E 238

Query 241 PKIHRMGDSFAVEIKGRHFDLFPVTRATNLPLTYLEAVYEAVLGTKTSKLGAAEIAAI 300
 Database 239 PKMQRGDMTAVEVGRHFDLHPLXHTTRTNLPLTYLEAVYEAVLGTKTSKLGAAEIAAI 300

Query 301 WETEESMKKLAQYSKEDARATYELGKEFFPMEEAELAKLIGQSYMDVSSSSTGNLVEWYLL 360
 Database 299 WESGENTERVAKYSMEDAKATYELGKEFFPMEEAELAKLIGQSYMDVSSSSTGNLVEWYLL 358

Query 361 RVAYEANELAQNPKDDEEYERRRLRITYLGGYKEBERGLWENITYDPLCYSITIWHN 420
 Database 359 RKAYEAEVANPKSSEYORRLRSYTGCFVKEBERGLWENITYDPLCYSITIWHN 418

Query 421 VSPDTLREGEKKNYDAPIVGKFCDFPFPISLGELEITMREIJKKMKATIDPIKK 480
 Database 419 VSPDTLNLEGCKNYDAPQVSHKFCDKDIPGPIPSLGHILLEERKIKTKMKETQDPIKKI 478

Query 481 MLDYRDAVLUHANSYGYGYPRARWYSKCEASVNTANGRYHETMIGEIEKEFGFVTL 540
 Database 479 LLDYRQKAIKULLANSEYGYGYAKARWYCEKAESVNTANGRYTIELVWKELEKFGFVTL 538

Query 541 YADTDGTYATIPIGEKPTIEKAKETLYKINSKLGLEIYEFGYLGREFVAKRYAVI 600
 Database 539 YIDTDGJYATIPIGEKPTIEKAKETLYKINSKLGLEIYEFGYLGREFVAKRYAVI 598

Query 601 DEEGRITRGLVEVRWWSAETOKAVLAIKEDSVKEAELVKDYYEEIARYQVPL 660
 Database 599 DEEGRITRGLVEVRWWSAETOKAVLAIKEDSVKEAELVKDYYEEIARYQVPL 658

Query 661 EKLVTHQITKDLSEYKAIGPHVATIARKLAAKGKIKVPRGTTISYIYLRSGSKISDRVIL 720
 Database 659 EKLVTHQITKDLSEYKAIGPHVATIARKLAAKGKIKVPRGTTISYIYLRSGSKISDRVIL 718

Query 721 SEDYDPKHKYDDYIENQNLPAVILAEAGYRKEDLYKQSSKQYGDALI 772
 Database 719 EYDPKHKYDDYIENQNLPAVILAEAGYRKEDLYKQSSKQYGDALI 770

RESULT 613
 US-08-688-649-37
 ; Sequence 37, Application US/08688649
 ; Patent No. 5,827,716
 ; GENERAL INFORMATION:
 ; APPLICANT: MAMONE, JOSEPH A.
 ; TITLE OF INVENTION: MODIFIED POL-II TYPE DNA
 ; NUMBER OF SEQUENCES: 38
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Lyon & Lyon
 ; STREET: 633 West Fifth Street
 ; CITY: Los Angeles
 ; STATE: California
 ; COUNTRY: U. S. A.
 ; ZIP: 90071-2066
 ; COMPUTER READABLE FORM:

Qy 601 DEEGRITRGLLEYVRDPWSEIAKETOAKVLEALKEDSVEKAIVEIYKDDVVBEIAKYQVPL 660
 Db 599 DEEGRVTRGLLEYVRDPWSEIAKETOAKVLEALKEDSVEKAIVEIYKDDVVBEIAKYQVPL 658
 Qy 651 EKLVTHEOITKOLSEY&KIGPHVIAKLAAGKIKYRPTLISIVYLRGSKISDRVLL 720
 Db 659 EKLAIEQTRPHEYKAGPHVAVAKLAAGKVKIKPMTGTYVLRGDGPISNRLLA 718
 Qy 721 SEYDPKCKYKDPPYYIENQVLPAVLRLAEGYRKEDLKYQSKQVGLAWL 772
 Db 719 EBYDPKCKYDABYYIENQVLPAVLRLAEGYRKEDLRYQXTRQVGLTSML 770

RESULT 14
 US-09-385-858-35
 i Sequence 35, Application US/09585858
 i Patent No 649161
 GENERAL INFORMATION
 i APPLICANT: Sigríður Hjörleifsdóttir
 i APPLICANT: Guðmundur O. Hrggviðasson
 i APPLICANT: Arnþor Ævarasson
 i APPLICANT: Jakob K. Kristjánsson
 i TITLE OF INVENTION: Bacteriophage RM378 of a Thermophilic
 i FILE REFERENCE: 2739.1001-001
 CURRENT FILING DATE: 2000-12-18
 PRIORITY APPLICATION NUMBER: 60/137,120
 PRIORITY FILING DATE: 1999-06-02
 NUMBER OF SEQ ID NOS: 73
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO: 35
 LENGTH: 773
 TYPE: PRT
 i ORGANISM: Desulfotoccoccus strain Tok
 US-09-585-858-35

Query Match 78.0%; Score 3141.5; DB 4; Length 773;
 Best Local Similarity 74.6%; Pred. No. 4e-236;
 Matches 577; Conservation 97; Mismatches 96; Indels 3; Gaps 2;

Qy 1 MIFDTDTYTKDKPIRKFKENGSEKIELDPHQDYIYALLKQDSDAETKAIGERG 60
 Db 1 MILDADTTEDDKPVTRFKERFKDPRDFEYIYALLKQDSDAETPIKITAERG 60
 Qy 61 KIVRVVDAVKVKKFLGRDVEWKLJIFEPDVPAIRGKIREHPAVIDYEYDIPFAKY 120
 Db 61 TTVRVTPAERVKKKFLGRPVEWKLJYFTHPDVPAIRDKIREHPAVIDYEYDIPFAKY 120
 Qy 121 LIDKGJLIPMEGDEELKLMADFETFYHEGDFGKEMIIMSYADEEARYTWNQDLY 180
 Db 121 LIDRGJLIPMEGDEELMLADFETLYHEGEGFGLM-SYADEEARYTWNQDLY 180
 Qy 181 VDVVSNREMIKRFVQIYREKDPDVLYTNGDNFDPYLLKRAEKGTYLIGRDKKEHPE 240
 Db 181 VESVSTKEMIKRFVQIYDPAFVLLKRSRNLGVFPLGDS-E 238
 Qy 241 PKITRMDSPAVEIKGRHDFLPYVTRTINLPYTLAEYAVLGIKTSKLGAEEIAI 300
 Db 239 PKITRMDSPAVEIKGRHDFLPYVTRTINLPYTLAEYAVLGIKTSKLGAEEIAI 298
 Qy 301 WETEESMKQAOYSMEDARAYELGKEFPMEEAELKLIGOSVWDSRSTGNLVEWLL 360
 Db 299 WESGEGLERVARYSMEDAKAYELGKEFPMEEAELKLIGOSVWDSRSTGNLVEWLL 358
 Qy 361 RVAYERNELAPNKPDEEYRRRLRTYLGIVKBERGWLWNTYDFRCLYPSIIVTHN 420
 Db 359 RKAYERNELVAPNKPDPERLARTE-SYAGGIVKBERGWLWNTYDFRCLYPSIIVTHN 417

Qy 421 VSPDTLREGCRNRYDAPIVYKFCKDFFPGRFPISTLGEITMROEIKKMKATIDPTEK 480
 Db 418 VSPDTLNREGREYDAPQVGHRCFDFFPGRFPISTLGEITMROEIKKMKATIDPTEK 477

Qy 481 MLDYQRATKLHANSYYGMYKARAVTSKECAESTWAGRAYITEMITKEIBKGRPKVL 540
 Db 478 LLDYQRATKLHANSYYGMYKARAVTSKECAESTWAGRAYITEMITKEIBKGRPKVL 537
 Qy 541 YADTDGFYATIPGKPKETIKKAKAFKLYINSKLPLGULEVEGFYLRLGFYAKORYAVI 600
 Db 538 YADTDGFYATIPGKPKETIKKAKAFKLYINSKLPLGULEVEGFYLRLGFYAKORYAVI 597
 Qy 601 DEEGRITRGLLEYVRDSEIAKETOAKVLEALKEDSVEKAIVEIYDVEEIAKYQVPL 660
 Db 598 DEEGRITRGLLEYVRDSEIAKETOAKVLEALKEDSVEKAIVEIYDVEEIAKYQVPL 657
 Qy 661 EKLVIEQTKDSEYKAGPHVAVLAARGKIKRPTGTLISIVYLRGSKISDRVLL 720
 Db 658 EKLVIEQTKDLSRATGPHVAVLAARGKIKRPTGTLISIVYLRGSKISDRVLL 717
 Qy 721 SEYDPKCKYDPPYYIENQVLPAVLRLAEGYRKEDLKYQSKQVLDAMW 773
 Db 718 DEEPAKERYDAAVYIENQVLPAVLRLAEGYRKEDLRYQTKQSLGMLK 770

RESULT 15
 US-08-375-134-12
 i Sequence 12, Application US/08375134
 i Patent No. 5602011
 GENERAL INFORMATION:
 i APPLICANT: Luhm, Robert A.
 i APPLICANT: d'Heanezel, Olga B.
 i APPLICANT: Duffaud, Guy D.
 i APPLICANT: Jolly, James F.
 i APPLICANT: Kelly, Robert M.
 i APPLICANT: Ting, Eve Yi-Tay
 i TITLE OF INVENTION: PURIFIED THERMOCOCCUS BAROSSI DNA POLYMERASE
 i NUMBER OF SEQUENCES: 13
 i CORRESPONDENCE ADDRESS:
 i ADDRESSEE: Charles & Brady
 i STREET: 411 East Wisconsin Avenue
 i CITY: Milwaukee
 i STATE: Wisconsin
 i COUNTRY: U.S.A.
 i ZIP: 53202-4497
 COMPUTER READABLE FORM:
 i COMPUTER: IBM PC compatible
 i OPERATING SYSTEM: PC-DOS/MS-DOS
 i SOFTWARE: Patent In Release #1.0, Version #1.25
 i CURRENT APPLICATION NUMBER: US/08/375,134
 i APPLICATION NUMBER: US/08/375,134
 i FILING DATE:
 i CLASSIFICATION: 415
 i ATTORNEY/AGENT INFORMATION:
 i NAME: Baker, Jean C.
 i REGISTRATION NUMBER: 35,433
 i REFERENCE/DOCKET NUMBER: 740211, 90440
 i TELECOMMUNICATION INFORMATION:
 i TELEFAX: (414) 277-5709
 i INFORMATION FOR SEQ ID NO: 12:
 i SEQUENCE CHARACTERISTICS:
 i LENGTH: 779 amino acids
 i TYPE: amino acid
 i STRANDEDNESS: single
 i TOPOLOGY: linear
 i MOLECULE TYPE: protein
 US-08-375-134-12

Query Match 78.0%; Score 3139; DB 1; Length 779;
 Best Local Similarity 76.2%; Pred. No. 6.4e-236;
 Matches 589; Conservative 81; Mismatches 99; Indels 4; Gaps 3;

Qy 1 MIFDTDTYTKDKPIRKFKENGFRXELDPHQYIYALLKQDSDAETKAIGERG 480

Db 1 MILDVYDITEDGKPVIRVFKVKDFGEFKIYDREEPYIYALLRDSAAEELEKITAERHG 60
 Qy 61 KIVRIVDVAVKVKCFPLGRDVEWKL1.FIEHQDFPALAR.-SKIREHPAVIDYEDIPFAKR 119
 Db 61 KIVRIVDVAVKVKCFPLGRSVEWVLYFTHQPAIRDPKIRGPAVIDYEDIPFAKR 120
 Qy 120 YLDIGLIPMEGDEBLKLKADEFIYHGDDEFEGEITIMIYADEBARVITWNKDLP 179
 Db 121 YLDIGLIPMEGDEBLKLKSFDELYHGEFTGTPLIMIYADEBARVITWNKDLP 180
 Qy 180 YDVYNSNERMIKREVOIYREKDPVLYTNGNDFDPLYKRAEKQGYTLIGRDKEP 239
 Db 181 YDVVSTEREMIKRJLKVKEFDPLVLYTGDNFDFAYLKRCKLQVSFTLGRDSS-- 238
 Qy 240 EPKIHRMGDSFAETKGRITHFDLFPVYRRTINLPYTLEAVYBVLGKTKSKIGAEETAA 299
 Db 239 EPKIHRMGDSFAETKGRITHFDLFPYTLEAVYBVLGKTKSKIGAEETAA 298
 Qy 300 IWESESMLNQYSMEDARATYELGEKFPMEELAKLQGSIWDFDVSSTONLWYL 359
 Db 299 AWETEGLEGARYSMEDARVYELGREFPMPMEQSLRJGQMDVSRSTGNLYWFL 358
 Qy 360 LRVAYERNEALPKDDEEVRRRTTYYGGYKEPERGLWNTYLDERCLYPSITVTH 419
 Db 359 LRVAYERNEALPKDDEEVRRRTTYYGGYKEPERGLWNTYLDERCLYPSITVTH 417
 Qy 420 NVSPDTLREBCKNYDVAPIVGYKFCDFPGFISIOLGELITMREQIKRMATIDPEK 479
 Db 418 NVSPDTLNREGCKSDVAPOVGHRECKDFPGFISIOLGELITMREQIKRMATIDPEL 477
 Qy 480 RMLDYZRQAVLHANSYYGTYNGYKARWYSKEACESVTAWGRHYTEMTEKEBEFKGKV 539
 Db 478 KLLDRQRATKILANSFGYGYARAWTCKEAEVTAWGRYTEMTRLESEKFGKD 537
 Qy 540 LYADTDFGYATIPEGKPEKIKKAKFELKYINSKLPGBLEEGFYLRGFFAKKRYAV 599
 Db 538 LYADTFLHATIPGADREYKQDLEFLNVTNPKLPGBLEEGFYSRGFFVTKKAV 597
 Qy 600 IDEEGRITTRGLEYVRDWESEIAKTOAKYLEATLKEDSYEKADEVIVDVEETAKYCP 659
 Db 598 IDEEGRITTRGLEYVRDWESEIAKETLARYLEATLRHGDEEATRIVVBETEKLSKYCP 657
 Qy 660 LEKLVHETOITKDISEYKAIGPHVIAKRLAAGJKVDRPOTISYVLRGSKKISDRVIL 719
 Db 658 PEKLVITEQITRELYKATGPHVIAKRAARGKIRPTEVSYVKGSGRLGDRAP 717
 Qy 720 LSEYDPKHKYDPDYYTENQVLPAVRLPAFGRKEDLKYQSSKQVGDawl 772
 Db 718 EDEFDPKHRDADYIENQVLPAVRLPAFGRKEDDERYQKTRQVGLGawl 770

Copyright (c) 1993 - 2004 GenCore version 5.1.6 CompuGen Ltd.

protein - protein search, using sw model

on: February 9, 2004, 11:46:55 [Search time: 40 Seconds
(without alignments)]
1860.864 Million cell updates/sec

searched: 283308 seqs, 96168682 residues

total number of hits satisfying chosen parameters: 283308

luminum DB seq length: 0
luminum DB seq length: 2000000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 Summary

abase : PIR 76:*

* 12

4: *Pir 4: **

No. is the number of results predicted by chance to have a score greater than or equal to the score of the total score distribution. This is derived by analysis of the total score distribution.

SUMMARIES

Query	Score	Match	Length	DB	ID
allnull	1	3356	83.4	1702	S42
allnull	2	3204	79.6	775	S42
allnull	3	3164	79.6	771	C75
allnull	4	3151	78.3	775	S35
allnull	5	2994	72.9	1312	S68
allnull	6	2886	72.9	1235	C71
allnull	7	2794	68.7	1670	S71
allnull	8	1301	32.3	1634	S64
allnull	9	1291	32.2	1634	S64
allnull	10	1205	30.0	586	C69
allnull	11	1204	29.9	581	C72
allnull	12	1190	29.6	803	B56
allnull	13	1138	28.3	784	E72
allnull	14	693	17.2	1088	T05
allnull	15	676	16.8	764	S75
allnull	16	659	16.4	1038	JCS
allnull	17	657	16.3	1038	T18
allnull	18	649	16.3	1086	T40
allnull	19	649	16.1	1086	T43
allnull	20	626	15.5	901	E84
allnull	21	621	15.4	1097	RNB
allnull	22	619	15.4	882	S23
allnull	23	616	15.4	1084	S19
allnull	24	616	15.3	872	JCT
allnull	25	616	15.3	882	F90
allnull	26	615	15.3	1094	S20
allnull	27	614	15.3	1094	S22
allnull	28	609	15.1	1107	A41
allnull	29	608	15.1	1107	ms2

30	605.5	15.0	1106	1	A39299	DNA-directed	DNA P
	587	14.6	1492	2	T18560	DNA-directed	DNA P
	584.5	14.5	879	2	A56277	DNA-directed	DNA P
	533.32	14.5	1462	1	DUHUA	DNA-directed	DNA P
	533.33	14.5	1081	2	T20698	hypothetical	protein
	581	14.4	875	2	JCS186	DNA-directed	DNA P
	578.5	14.4	875	2		probable	DNA-directed
	570.5	14.2	959	2	T27663	DNA-directed	DNA P
	570	14.2	1465	2	S456228	DNA-dependent	DNA P
	559.5	13.9	223	2	S69125	DNA-directed	DNA P
	548.5	13.6	1339	1	S20052	DNA-directed	DNA P
	527.5	13.1	1468	2	S58250	DNA-directed	DNA P
	522.5	13.0	1505	2	S28079	DNA-directed	DNA P
	516.5	12.8	844	2	T31321	DNA-directed	DNA P
	516	12.8	1405	1	DZPZA	hypothetical	protein
	504.5	12.4	2154	2	A846679	DNA-directed	DNA P
	500.5	12.4	913	2	T17675	DNA-directed	DNA P

AGREEMENTS

RESULT 1

lining precursor - Thermococcus
Release (EC 3.1.-.-) PI-II, DR
Change 18-Jun-1999
B.:i Kucera, R.B.; Benner, J.

•1; PID:9154686
B.; Kucera, R.B.;
gene.

1099;1466-1489;1533-1547 <PE>
an archaea DNA polymerase.

8

Db	1 MILDITYITKDGKPITRIFKENGEEFKIEDLDPHQPYTALLKDSAEEBIAKTGEEHG	60	Db	1081 DSVGESEIIIRONGKIRFYKIKDQLFSKVDYSIGBKEYCITLEGYEALTLDDGKLWNPV	1140	
Qy	61 KIVRYVDYAVKVKKKFLGRDVEWMKLFEPDQFALKIREHPDQVYDIPFARY	120	Qy	544	-----	
Db	61 KIVRYVDYAVKVKKKFLGRDVEWMKLFEPDQFALKIREHPDQVYDIPFARY	120	Db	1141 PYVMRHRANKMFRWTWNTSMWYDTEHDHSIUGVNTSKTKTAKIGERLKEVKPFELGK	1200	
Qy	121 LIDKGLIPMEGDEELKLMAFDIETFYHGDEFKGKBEIMISYADEEAVITWKNIDLPY	180	Qy	544	-----	
Db	121 LIDKGLIPMEGDEELKLMAFDIETFYHGDEFKGKBEIMISYADEEAVITWKNIDLPY	180	Db	1201 AVKSLICPNAPLKDENTKTSEIAVKFWELVGLIVSGDGNGGDSRWAEYIYLGLSTGKDAAE	1260	
Qy	181 VDVSNEREMKRFQIVREKDPYLIITYGDNFLPILYIKRAEKLGLVTLIGRDKEPE	240	Qy	544	-----	
Db	181 VDVSNEREMKRFQIVREKDPYLIITYGDNFLPILYIKRAEKLGLVTLIGRDKEPE	240	Db	1261 IKQKLLBLPLKTYGVISNYYPKNEKGDFENILAKSLIKFEMKRFHDKGKRRKIPETMYELPV	1320	
Qy	241 PKIHRNGDSFAVEIKRGIHFOLFPYVTRTINLPYTLTLEAVEAVLGKTKSKLGAEELAAI	300	Qy	544	-----	
Db	241 PKIQRMDSPFAVEIKRGIHFOLFPYVTRTINLPYTLTLEAVEAVLGKTKSKLGAEELAAI	300	Db	1321 TYIEAFDRLGLFSADGTVTIRKGVPETRLTNIDADFLIREVRLKLWVGIINSIPIATTPNR	1380	
Qy	301 WETEESMKLQYSMEDARATYELGKFFPMEAEALAKLGQSVDVSRSSTGNLVEWIL	360	Qy	544	-----	
Db	301 WETEESMKLQYSMEDARATYELGKFFPMEAEALAKLGQSVDVSRSSTGNLVEWIL	360	Db	1381 YNGVSTGTYSKHLIRKNTKWRFAERIGFLIERKQRKLELEHKSARVTRNTIDFGFLVHVK	1440	
Qy	361 RAYERELAPKPDDEBEYERLRLTLYGGYKEPERGLMENITYDFRLGLYPSITVHN	420	Qy	544	-----	
Db	361 RAYERELAPKPDDEBEYERLRLTLYGGYKEPERGLMENITYDFRLGLYPSITVHN	420	Db	1441 KVEEIPYEGYYDIEVEETHFFANNLVLTNTDGFATIPEKPLIKCCKAKEKFNYINS	1500	
Qy	421 VSPDTLFRGCKKNYDAPIVGKPKCDFPGFTPSLIGELTMROBIKGMKATIDPIECK	480	Qy	573 KLPGLLLEYEFSYLRGFFYAKRVAVIDEGRGTRTGLEVRRDSEIAKETOAKVLEA	632	
Db	421 VSPDTLFRGCKKNYDAPIVGKPKCDFPGFTPSLIGELTMROBIKGMKATIDPIECK	480	Db	1501 KLPGLLLEYEFSYLRGFFYAKRVAVIDEGRGTRTGLEVRRDSEIAKETOAKVLEA	1560	
Qy	481 MLDYRQRAVKLHA-----	493	Qy	633 ILKEDSYEKALEVQDVEELAKYQVPLKVEIHEQITKDLSEYKXAGIPHYAIAKRLAAK	692	
Db	481 MLDYRQRAVKLANSILIPNEWLPIIENGELIKEVKVKGAYEIQLOSSGRKINTAGHSLFTVRNGEI	540	Db	1561 ILKEDSYEKALEVQDVEELAKYQVPLKVEIHEQITKDLKXAGIPHYAIAKRLAA	1620	
Qy	494 -----	493	Qy	693 GIKVRPGETIISITVLRGSGKISDRVILLSEYDPKHYDPPYIENQVLPAVLRLBAFG	752	
Db	541 VNNLFAFSFNKKIKESFVKVKAALIRHKYKGKAYEIQLOSSGRKINTAGHSLFTVRNGEI	600	Db	1621 GIKVPGIIISITVLRGSGKISDRVILLSEYDPKHYDPPYIENQVLPAVLRLBAFG	1680	
Qy	494 -----	493	Qy	753 YRKEDLKYQSSKOVGDAWLK	774	
Db	601 KEVSGDGKIGEKGDLIVPKKIKLNEKGYSINIPELISDSEETADIVMTISAKGRKNEFK	660	Db	1681 YRKEDLRYQSSKOTGDAWLKR	1702	
Qy	494 -----	493	RESULT 2			
Db	661 GMRLTRLAMFGEENRRIRTNFLYFLHKLGLIUKLRLRGYEVTDNERLKKYKOLYELAG	720	S67220			
Qy	494 -----	493	DNA-directed DNA polymerase (EC 2.7.7.7) - Thermococcus sp.			
Db	721 SVKYNGKREYLWFMNIEIKDFISYPPKKELEWKIGTLNGFRNTCLKVDEDFGKLGY	780	C:Species: Thermococcus sp.			
Qy	494 -----	493	C:Accession: S67220			
Db	781 VSEGYAGAQNKNTGGIISYSVSKYNEDDNVLESMKVNAKEFKGKVRDVRNCVSISSKMMAYL	840	C:Date: 17-Feb-1998 #sequence_revision 13-Mar-1998			
Qy	494 -----	493	C:Accession: S67220			
Db	901 VFLINSLGTSVSKIGFSGVYRVYINBDQFFPQTSSRKNTYYSNLPIKEILRDVFKEFQ	960	C:Submitted to the EMBL Data Library, January 1996			
Qy	494 -----	493	C:Description: Cloning, expression and modulation of the 3'-5' exonuclease activity of			
Db	961 KMTFPKFKELVDSGKLNREREAKLLEFFINGDIVDRVKSYKEKDYEYYDLSVEDNEN	1020	C:Accession: S67220			
Qy	494 -----	493	A:Residues: 1-775 <SOU>			
Db	841 VMKCLCGALAENRIPSVLTSPEPVRSWSFLEAYFTGDIDHPSKRFLSTKSELANQL	900	A:Cross-reference: GB:U47108; NID:91197451; PID:AAA88769.1; PID:91197452			
Qy	494 -----	493	A:Experimental source: strain 90-7			
Db	961 KMTFPKFKELVDSGKLNREREAKLLEFFINGDIVDRVKSYKEKDYEYYDLSVEDNEN	1020	C:Superfamily: herpesvirus DNA-directed DNA polymerase			
Qy	494 -----	493	C:Keywords: DNA binding; nucleotidyltransferase			
Db	901 VFLINSLGTSVSKIGFSGVYRVYINBDQFFPQTSSRKNTYYSNLPIKEILRDVFKEFQ	960	Qy	1 MIFPTDYLTKDGKPKIIRIFKENGEGFKFIELDPHQPYTIALLKDSAIDSEKAKGERH	60	
Qy	494 -----	493	Db	1 MIFPTDYLTKDGKPKIIRIFKENGEGFKFIELDPHQPYTIALLKDSAIDSEKAKGERH	60	
Db	841 VMKCLCGALAENRIPSVLTSPEPVRSWSFLEAYFTGDIDHPSKRFLSTKSELANQL	900	Qy	61 KIVYVVDAYKVKKCFGLGRDVEWMLKIFEPDQFALKIREHPAYDIDYDIPFAKRY	120	
Qy	494 -----	493	Db	61 TVVKVRAEKVQKRFGRPIEWKLYNHFQDPAIRDRIRAHAVVYDIPFAKRY	120	
Qy	543 D-----	543	Qy	61 TVVKVRAEKVQKRFGRPIEWKLYNHFQDPAIRDRIRAHAVVYDIPFAKRY	120	

Qy	121 LIDKGLIPMEGDBELKLMAPDILETFHDEEFGKELIMISYADEEPRAVITWNIDPY 180	Qy	61 KIVRVVDAVKVKKKFLGRDVEWKLIFEPHQDVPALRGKIREHPAVIDYEDIFPAKRY 120
Db	121 LIDKGLIPMEGDBELKLMAPDILETFHDEEFGKELIMISYADEEPRAVITWNIDPY 180	Db	61 KIVRTEVEVKVKKKFLGRDVEWKLIFEPHQDVPALRGKIREHPAVIDYEDIFPAKRY 120
Qy	181 VDYSNEEREMIKRFVQTYREKPDVLLTYNGNFDPLYLRAEKGVTLILGRDEKHEPE 240	Qy	121 LIDKGLIPMEGDBELKLMAPDILETFHDEEFGKELIMISYADEEPRAVITWNIDPY 180
Db	181 VDYSNEEREMIKRFVQTYREKPDVLLTYNGNFDPLYLRAEKGVTLILGRDEKHEPE 240	Db	121 LIDKGLTPMEGNEBLTFLAVDILTYHEGEERFKGKPTIMISTADEEGAVITWKSIDPY 180
Qy	241 PKHRMGFSAVEIKGRIHDFLPVWRITINLPPTYLEAVVAKTTSKSLGABETAAI 300	Qy	181 DVYVSNEREMIKRFVQTYREKPDVLLTYNGNFDPLYLRAEKGVTLILGRDEKHEPE 240
Db	239 PKTQRMGRFAVEVKGRIHDFLPVWRITINLPPTYLEAVVAKTTSKSLGABETAAI 298	Db	181 DVYVSNEREMIKRFVQTYREKPDVLLTYNGNFDPLYLRAEKGVTLILGRDEKHEPE 240
Qy	301 WETBESEMVKLAQYSMEDARATYELGEFFPMDEAELAKLIGOSWDVRSSTGNLYEWYL 360	Qy	241 PKHRMGDSFAVEIKGRIHDFLPVWRITINLPPTYLEAVVAKTTSKSLGABETAAI 300
Db	299 WEGEGLEVRARYSMEDAKVYELVGRGFFPMDEAQLSLIGOSLWDVRSSTGNLYEWYL 358	Db	239 PKHRMGDSLAVEIKGRIHDFLPVWRITINLPPTYLEAVVAKTTSKSLGABETAAI 298
Qy	361 RYAYERNLAPKLPKDEEYRRLRTTLYGGYKEPERGLWENITYLDFRCLYPSLIVTHN 420	Qy	301 WETBESEMVKLAQYSMEDARATYELGEFFPMDEAELAKLIGOSWDVRSSTGNLYEWYL 360
Db	359 RKEYKRNELAPKLPKDERLLARR-RGGAGGYKEPERGLWENITYLDFRCLYPSLIVTHN 417	Db	299 WETBEGLERVKAKYSMEDAKVYELVGRGFFPMDEAQLSLIGOSLWDVRSSTGNLYEWYL 358
Qy	421 VSPDTLREREGCKNQDYAFIGYKPCFKCDFPGFIPSILGELTMROEIKKRMKATIDPKKK 480	Qy	361 RYAYERNLAPKLPKDEEYRRLRTTLYGGYKEPERGLWENITYLDFRCLYPSLIVTHN 420
Db	41.8 VSPDTLREREGCKNQDYAFIGYKPCFKCDFPGFIPSILGELTMROEIKKRMKATIDPKKK 477	Db	359 RYAYERNLAPKLPKDEEYRRLRTTLYGGYKEPERGLWENITYLDFRCLYPSLIVTHN 418
Qy	481 MLDYRQZAVKLANSYSSYMGYPKARYSKCAGESTAWAGHYIETIKELEEKFGKVL 540	Qy	421 VSPDLEREGCKNQDYAFIGYKPCFKCDFPGFIPSILGELTMROEIKKRMKATIDPKKK 480
Db	478 LDYRQZAVKLANSYSSYMGYPKARYSKCAGESTAWAGHYIETIKELEEKFGKVL 537	Db	419 VSPDLERENCKEYDVAQVGHFPCFKCDFPGFIPSILGELTMROEIKKRMKESKDPTBK 478
Qy	541 YADTDGFYATIPEKEPKTIIKKAKEFLKYINSKLPGILEEYGFYLRGFIVAKRYAVI 600	Qy	481 MLDYRQAVKLHANSYGYMGYPKARYSKCAGESTAWAGHYIETIKELEEKFGKVL 540
Db	538 YADTDGFYATIPEKEPKTIIKKAKEFLKYINSKLPGILEEYGFYLRGFIVAKRYAVI 597	Db	479 LDYRQZAVKLANSYSSYMGYPKARYSKCAGESTAWAGHYIETIKELEEKFGKVL 537
Qy	601 DEGRITTRGLEVRRDWEIAKEQTOQAVLEALKEDSVERAKEYVYDQWEIARYQVPL 660	Qy	541 YADTDGFYATIPEKEPKTIIKKAKEFLKYINSKLPGILEEYGFYLRGFIVAKRYAVI 600
Db	598 DEEGKTRITRGLEVRRDWEIAKEQTOQAVLEALKEDSVERAKEYVYDQWEIARYQVPL 657	Db	538 YADTDGFYATIPEKEPKTIIKKAKEFLKYINSKLPGILEEYGFYLRGFIVAKRYAVI 597
Qy	661 EKVTHEQDITKODSEYKAGPVAIAKRLAAGIKYRPTGTTISYIYLRSKSKISDRVIL 720	Qy	601 DEGRITTRGLEVRRDWEIAKEQTOQAVLEALKEDSVERAKEYVYDQWEIARYQVPL 660
Db	658 EKVTHEQDITKODSEYKAGPVAIAKRLAAGIKYRPTGTTISYIYLRSKSKISDRVIL 717	Db	598 DEEGKTRITRGLEVRRDWEIAKEQTOQAVLEALKEDSVERAKEYVYDQWEIARYQVPL 657
Qy	721 SEYDPKCKHYDPPYTTENQVLPAVLRLTEAFLGSKYRKEDELYKQYSSKQVGLDAWLK 773	Qy	661 EKVTHEQDITKODSEYKAGPVAIAKRLAAGIKYRPTGTTISYIYLRSKSKISDRVIL 720
Db	718 DEFPTRFHDYDYYIENQVLPAVLRLTEAFLGSKYRKEDELYKQYSSKQVGLDAWLK 770	Db	658 EKVTHEQDITKODSEYKAGPVAIAKRLAAGIKYRPTGTTISYIYLRSKSKISDRVIL 717
RESULT 3			
C75023	DNA Polymerase I PAB1128 - Pyrococcus abyssi (strain Orsay)	Qy	721 SEYDPKCKHYDPPYTTENQVLPAVLRLTEAFLGSKYRKEDELYKQYSSKQVGLDAWLK 773
C.Species: Pyrococcus abyssi		Db	718 EEFDPKCKHYDYYIENQVLPAVLRLTEAFLGSKYRKEDELYKQYSSKQVGLDAWLK 770
C.Date: 20-aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000			
R; anonymous Genoscope			
submitted to the EMBL Data Library, July 1999			
A; Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome structure			
A; Reference number: A75001			
A; Accession: C75023			
A; Status: preliminary			
A; Molecule type: DNA			
A; Residues: 1-771 <RAW>			
A; Cross-references: GB:AUJ248288; GB:AL096836; NID:95458960; PID:CAB50625.1; PMID:93181200; MUID:8441634			
A; Experimental source: strain Orsay			
C; Genetics: PAB1128			
C; Superfamily: herpesvirus DNA-directed DNA polymerase			
Query Match 78.6%; Score 3164.5%; DB 2; Length 771;			
Best Local Similarity 76.1%; Pred. No. 4.6e-155; Indels 3; Gaps 2;			
Matches 588; Conservative 89; Mismatches 93;			
Qy 1 MIFDTDVTIKDGKPIIRFKENGEFKIYDHFQPYTIVALLKDDSAIDELKAIKGERIG 60			
Db 1 MIDADYTEDSKPIIRFKENGEFKIYDHFQPYTIVALLKDDSAIDELKAIKGERIG 60			
RESULT 4			
S35543	DNA-directed DNA polymerase (BC 2.7.7.7) - Pyrococcus furiosus	C;Species: Pyrococcus furiosus	
C;Accession: C75023		C;Date: 31-Dec-1993 #sequence_revision 02-Aug-1994 #text_change 20-Jun-2000	
A;Description: S35543; S44596		C;Accession: S35543; S44596	
R;Demori, T.; Ishino, Y.; Toh, H.; Asada, K.; Kato, I.		R;Demori, T.; Ishino, Y.; Toh, H.; Asada, K.; Kato, I.	
Nucleic Acids Res 21, 259-265, 1993		Nucleic Acids Res 21, 259-265, 1993	
A;Title: Organization and nucleotide sequence of the DNA polymerase gene from the archa		A;Title: Organization and nucleotide sequence of the DNA polymerase gene from the archa	
A;Reference number: S35543; MUID:93181200; PMID:8441634		A;Reference number: S35543; MUID:93181200; PMID:8441634	
A;Molecule type: DNA		A;Molecule type: DNA	
A;Residues: 1-775 <UEM1>		A;Residues: 1-775 <UEM1>	
A;Cross-References: EMBL: D12983; NID:9216917; PID: BAA02362.1; PMID:9216918		A;Cross-References: EMBL: D12983; NID:9216917; PID: BAA02362.1; PMID:9216918	
A;Accession: S44596		A;Accession: S44596	
A;Molecule type: protein		A;Molecule type: protein	
A;Residues: 1-12 <UEM2>		A;Residues: 1-12 <UEM2>	
A;Gene: polI		A;Gene: polI	
C;Superfamily: herpesvirus DNA-directed DNA polymerase		C;Superfamily: herpesvirus DNA-directed DNA polymerase	
E;1-775/Product: DNA-directed DNA polymerase		E;1-775/Product: DNA-directed DNA polymerase	
Query Match 78.3%; Score 3151; DB 2; Length 775;			

Best Local Similarity 74.6%; Pred. No. 2.3e-154;	
Matches 576; Conservative 100; Mismatches 94; Indels 2; Gaps 1;	
Qy	1 MIFPDYITKDGPIPIIRFCKENGFRKIELDPFPQTYIALLKDSAIDBIKAKGEBHG 60
Db	1 MILEDVDTYIEEGPKVIRLKKENGKFKIEHDTTRPYIALLRDSKTEVKKTGERG 60
Qy	61 KIVRVDVAVKVKCKFLGRDVEYKLLIEHFPQDVPALRGKIREHPAVDIFYDIPFARRY 120
Db	61 KIVIIVDVEKVKFLGPITWKLIEHFPQDVPTRKETREHPAVDIFYDIPFARRY 120
Qy	121 LIDGLIPMEGDFELKUMAFDIEFYTHEGDFGKGBLIMISYADEBEEARYITWNIDLY 180
Db	121 LIDGLIPMEGDFELKIAFD-ETLYHGEETGKPPIMISYADEEAKYITWNIDLY 180
Qy	181 VDVNSNEREMIKEFVQIVREKDPDVLIYTNGNFDLPLYKRAEKLGVTLLGRKEHPE 240
Db	181 VEVVSNEREMIKEFVQIVREKDDIVTVNGNSPDDFLYKRAEKLGKATIGRGS-E 238
Qy	241 PKHIRMGDSFAVEIKGRTHFDLPPVVRTINLPTTYLEAVYEAVLGKTKSKLGAEBIAAI 300
Db	239 PKMQRIGDMTAVEKGRTHFDLYHVTCTINLPTTYLEAVYEAIFPKPKERVKYABDIARA 298
Qy	301 WETBESMRKLLAQYMEDARATYBKGKEFPMEALAKLIGQSVDWYDSSSTGCLNLYEWYIL 360
Db	299 WESGENLERVAKTSMEDAKATYEGKEELPMETQSLSLVQPLWYDVSRSSTGCLNLYEWYIL 358
Qy	361 RVAXERNELAQNPKDDEEYRRRLRTTYLGGYREPERGLWENITYDFRLCLYPSIVTHN 420
Db	359 RKAVERNEVAPNKEKSEEEYQRRRLRESEITGGFVKEPERGLWENIVYDFRAYPSLITHN 418
Qy	421 VSPDTLEREGCKNDVADPVGKCKDQDPGFTPSISGLBLTMHQEIKKMKATIDPIECK 480
Db	419 VSPDTLNLEGCKNDVADPQNGHKECKD1PGF1ISLGLHLLERQKTKMKETQDDIEKI 478
Qy	481 MLDYQROAVKLHANSYYGGMGYPKARYWSKECAESVATGHWYIEMTKEBEKGFKV1 540
Db	479 LLDYQROAVKLHANSYYGGMGYPKARYWSKECAESVATGKRY1ELWKEBEKGFKV1 538
Qy	541 YADTGCYATIPGKEPKETKKAKEBFLKYINSKLPGILEYEGFYLRGFPPVAKKRYAVI 600
Db	539 YTDTDDLYATPGGSEELKCKKALEFVKYINSKLPGILEYEGFYLRGFPPVAKKRYAVI 598
Qy	601 DEEGRGTTGLEYVIRDWSETAKTQAKYLAIIKEDSVKAVEIVDVEBIAKQVPL 660
Db	599 DEEGRKVITGLEYVIRDWSETAKTQAKVLETILKHGDOVEAVTRIVKQVLANEIEPP 658
Qy	661 EKLVTHEQITDLSSEYKAIGPHVATAKRLAAGKIKVPRGTSIISYIVLRSGSKISDRVIL 720
Db	659 EKLAVIEQITRPLHEKAIPHVAVAKLAAKGKIKVGMVIGYIVLRGDGTSINNAILA 718
Qy	721 SEYDPKHKYDPPDYYTENQVLPAVRLILEAFGTYKEDLKYQSSKQVGDawl 772
Db	719 E3YDPKHKYDAAEYIENQVLPAVRLILEFGTYKEDLRYQKTRQVGLTSKL 770

Qy	4 94	-----	493	Db	61	KVRII VETEKIQRKLLGRPTEWWKCLYLERPQDPAIRDKREHAAVD FEYDIPFAKRY 120
Db	719	VRYNGNKEYLVEFNADVVISLMPPEELKEWRIGTRNGFRMSTFVIDEDPAKLQYYV 778	Qy	121	LIDKG LI PMGDEELKLMAIDIEYTHEGDEFGKGEIIMSYADEBEEARVITWNIDLY 180	
Qy	4 94	-----	493	Db	121	LIDKG LI PMGDEELKLMAIDIEYTHEGDEFGKGEIIMSYADEBEEARVITWNIDLY 180
Db	779	SEGSARKWNQTGWSYITVLYNENDEVLDMEHLAKKFPGKVRGKNYVEIPKRYAYII 838	Qy	181	VDVVSNERENTIKRFQIVREKDPDVLITYNGDNDPLPYLIKRAEKGYTLLGRDKEHPE 240	
Qy	4 94	-----	493	Db	181	VEVVSNERENTIKRFQIVREKDPDVLITYNGDNDPLPYLIKRAEKGYTLLGRD--NSE 238
Db	839	FESLCGTLAENKRVPEVIFTSSKSVRWAFLLEGFIGDGVHPSKRVRLSTKSSELLYNGIV 898	Qy	241	PKHHRMGDSFAVETKGRHFDLFPYVVRTINLPYTLRVEAFLGKTKSGLAEELIAI 300	
Qy	4 94	-----	493	Db	239	PKMQMGDSLAVEKGRHFDLFPYVVRTINLPYTLRVEAFLGKTKSGLAEELIAI 298
Db	899	LLNSLGLGSAIKUGYDSSGTVYRVTYNEELURKFTERYRKCCNIVYSH1VPKDILKETFGKVFQK 958	Qy	301	WETEBSMKKLAQSYMEDARATYEGKEPPMEABJLQIGSQWDSSESTGNLVEWLL 360	
Qy	4 94	-----	493	Db	299	WETGGLERVAKYKSMEDAKTYELGREFFPMEAGLARLQVGPWDSRSSTGNLVEWLL 358
Db	959	NISYKKFRELVNGKLDREREKAKRRIEWLNLNGDIVLDRVYEIKREYDGYVYDLSVDEDNF 1018	Qy	361	RVAYERNELAPNKPDEEVYRRLPITYLGGYKEPERGLWENTYLDERCLYPSIIVTHN 420	
Qy	4 94	-----	493	Db	359	RKAYERNELAPNKPDEEVYRRLPITYLGGYKEPERGLWEGIVSLDPLRSVPSIITAN 418
Db	1019	LAGFGFLVYAHNSYSGYGGYAKARWYCKCAESUTAWGHYIEMTKEERKGFKVLYD 543	Qy	421	VSPDLEREECKNDVAPIVGYKECKDFPFGFIPSISIGELITYMROEIKERMRATIDIEKK 480	
Qy	544	TDGFYATPGKEKPTIKGKAEFLKYGINSKLPGSLELEYEGFYLRGFFVAKRYAYIDEE 603	Db	419	VSPDLINREGCBEVDAVPGHFRCKDTPGFPISLGLOLLEEROKIKCRMKESKDIPVKK 478	
Db	1079	TDGUYATPGAKPEEKKAKALEFVYDYNAKLPGSLELEYEGFYVRGFVTKRYALIDEE 1138	Qy	481	MLDYRQAVKL-----	
Qy	604	GRITTRGIEVVRDWSSEAKETOKAIVLKEAILEDSEVKADEVETAKYQVPLEXL 663	Db	479	LLDLQRKA KILANSILPDEWLPIVENEKVRFVKGDFIDREIBENAEVRKDGETEILE 538	
Db	1139	GK1ITRGIEVVRDWSSEAKETOKAIVLKEAILEDSEVKADEVETAKYQVPLEXL 1198	Qy	492	-----	
Qy	664	VTEHQITKDLSEYKAIKGPHVATKRLAAGKGIVKVRGCFIISYTVLRSGGKISDRVILSEY 723	Db	539	VKDMLALSFRRETKESELKXKVALIRHRSGKVYSIKLGSRRKITSGSHLSFSVNGKL 598	
Db	1139	VTEQITPLHETKAIGHVAVAKRLAARGVVRPGMVIGITVLRSGDPIKRATLAEF 1258	Qy	492	-----	
Qy	724	DPKKHYDPDYXLENQVLPAVLRLAEFGYRKEDLKYQSSKQVQGLDAWL 772	Db	599	VKVRGDELKPGDLYVVPGRKLPESKOVLNLYELLKLPEBETSNVMMIPYKGRKNFFK 658	
Db	1259	DLRKHYTDAYEYIENQVLPAVLRLAEFGYRKEDLRYQTKTGJTLAWL 1307	Qy	492	-----	
Db	-----	-----	Db	659	GMLKLTWYIFGEGEGERPTDAGRVLKHLERLGYYKLKRGCEVLDWEISLKRKLYLTKN 718	
Qy	492	-----	Qy	492	-----	
Db	719	LKYNGNSRAYMVFNSLRDVVSLSMPPEELKEWIGEPRGPKGTFIDVDDSPAKLGYI 778	Db	719	LKYNGNSRAYMVFNSLRDVVSLSMPPEELKEWIGEPRGPKGTFIDVDDSPAKLGYI 778	
Qy	4 92	-----	Qy	492	-----	
Db	779	SSGVEDKDRVKFHSKDQNVLEDIAKLAEKLFGKVRGRGYIEVSGKISHA1FRVLAEGKR 838	Db	779	SSGVEDKDRVKFHSKDQNVLEDIAKLAEKLFGKVRGRGYIEVSGKISHA1FRVLAEGKR 838	
Qy	4 94	-----	Qy	494	-----	
Db	839	IPEFFITSPMD1KVAFLGQNGNAEEDTFSTKSELLVNQLTLLNSIGVSDIKEBEKGV 898	Db	839	IPEFFITSPMD1KVAFLGQNGNAEEDTFSTKSELLVNQLTLLNSIGVSDIKEBEKGV 898	
Qy	494	-----	Qy	494	-----	
Db	899	YRYINKKESNSNDIVLDSVSEVEREGYTYDLSVENDNFYFLGFLLYAHNSSTYGY 958	Db	899	YRYINKKESNSNDIVLDSVSEVEREGYTYDLSVENDNFYFLGFLLYAHNSSTYGY 958	
Qy	501	GPKKARYSKKECAESTAWGRHKAEMTIKEEKKFVKVLADTOFVATIPGKR-ETI 559	Db	501	GPKKARYSKKECAESTAWGRHKAEMTIKEEKKFVKVLADTOFVATIPGKR-ETI 559	
Db	959	GTAKARYCKKECAESTAWGRHKAEMTIKEEKKFVKVLADTOFVATIPGKVWEV 1017	Db	959	GTAKARYCKKECAESTAWGRHKAEMTIKEEKKFVKVLADTOFVATIPGKVWEV 1017	
Qy	560	KKKAKEFILKVKINSKLPLLELEYEGYLRGPVAKRVADEEGRITTRGLEYVRDWS 619	Qy	560	KKKAKEFILKVKINSKLPLLELEYEGYLRGPVAKRVADEEGRITTRGLEYVRDWS 619	
Db	1018	XPKALEVYDYNSKLPCVLELEYEGYARGFVTKKYALDEEKKIVTRGLEYVRDWS 1077	Db	1018	XPKALEVYDYNSKLPCVLELEYEGYARGFVTKKYALDEEKKIVTRGLEYVRDWS 1077	
Qy	620	EIAKETOAKVLEAILEDSEVKAEVKDVYBEEIARYQVPLKLYTHEQITKDLSBEYKAI 679	Qy	620	EIAKETOAKVLEAILEDSEVKAEVKDVYBEEIARYQVPLKLYTHEQITKDLSBEYKAI 679	
Db	1078	EIAKETOAKVLEAILEDSEVKAEVKDVYBEEIARYQVPLKLYTHEQITKDLSBEYKAI 1137	Db	1078	EIAKETOAKVLEAILEDSEVKAEVKDVYBEEIARYQVPLKLYTHEQITKDLSBEYKAI 1137	
Qy	680	GPHVIAKRLAAGKIVKVRPGFLISYVLRGSGKSKDRVILSEYDPKKHYDPYVYIENQ 739	Qy	680	GPHVIAKRLAAGKIVKVRPGFLISYVLRGSGKSKDRVILSEYDPKKHYDPYVYIENQ 739	

RESULT 6

C71210 Probable DNA-directed DNA polymerase - Pyrococcus horikoshii
 C Species: Pyrococcus horikoshii
 C Date: 14-Aug-1998 #Sequence_revision 14-Aug-1998 #text_change 21-Jul-2000
 C Accession: C71210
 R: Kawarabayasi, Y.; Sawada, M.; Horikawa, H.; Hino, Y.; Yamamoto, S.; Sekine M.; Ohfuki, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kusuda, N.; Oguchi DNA Res. 5, 55-76, 1998
 A: Title: Complete sequence and gene organization of the genome of a hyper-thermophilic archaeon: *Archaeoglobus fulgidus* strain KF-7001
 A: Reference number: A71000; MUID:98344137; PMID:9679194
 A: Accession: C71210
 A: Status: Preliminary; nucleic acid sequence not shown; translation not shown
 A: Molecule type: DNA
 A: Residues: 1-1235 <KAW>
 A: Cross-references: GB:AP000007; NID:g3236134; PID:BAA31074.1; PID:g3255391
 A: Experimental source: strain OT3
 A: Note: this accession replaces an interim accession for a sequence replaced by GenBank
 C: Genetics:
 A: Gene: PH1947
 C: Superfamily: hypothetical protein PH0202

Query Match 574; Score 2886; DB 2; Length 1235;
 Best Local Similarity 46.5%; Pred. No. 1.7e-140;
 Mismatches 89; Indels 464; Gaps 5;
 A: Residues: 1-1235 <KAW>
 A: Cross-references: GB:AP000007; NID:g3236134; PID:BAA31074.1; PID:g3255391
 A: Note: this accession replaces an interim accession for a sequence replaced by GenBank
 C: Genetics:
 A: Gene: PH1947
 C: Superfamily: hypothetical protein PH0202

Query Match 574; Score 2886; DB 2; Length 1235;
 Best Local Similarity 46.5%; Pred. No. 1.7e-140;
 Mismatches 89; Indels 464; Gaps 5;
 A: Residues: 1-1235 <KAW>
 A: Cross-references: GB:AP000007; NID:g3236134; PID:BAA31074.1; PID:g3255391
 A: Note: this accession replaces an interim accession for a sequence replaced by GenBank
 C: Genetics:
 A: Gene: PH1947
 C: Superfamily: hypothetical protein PH0202

Query Match 574; Score 2886; DB 2; Length 1235;
 Best Local Similarity 46.5%; Pred. No. 1.7e-140;
 Mismatches 89; Indels 464; Gaps 5;

Db	1138 GPHVAVAKRLMARGIKYKPGMVTGYIVLRLGDPISKRAISIEEDPRKHYDAEYYIENQ 1197	Db	359 RKAVERNELAPNKDEKEKLARR - RQSVEGGYKKEPERGLWENIVYDFRCHPADTUVVK 417
Qy	740 VLPAVLRLDEAFSYRKEDLKYQSKQVGLDAWLUK 773	Qy	411 -----
Db	1198 VLPAAVERLILKAFSYKREDLRQKTKQVGLGANTK 1231	Db	410 GKGIIINISVEQEGDYVLGIDGMWORVVKWEDYDYGELVNINGLKCPTNHLKPVTVNERO 477
RESULT 7			
S71551	DNA-directed DNA polymerase (EC 2.7.7.7) KOD, intein containing precursor - Pyrococcus	Qy	411 -----
NC: Contains: DNA endonuclease (EC 3.1.-.-) PI-I; DNA endonuclease (EC 3.1.-.-) PI-II; DNA	C: Species: Pyrococcus sp.	Db	418 GKGIIINISVEQEGDYVLGIDGMWORVVKWEDYDYGELVNINGLKCPTNHLKPVTVNERO 477
A: Variety: strain KOD1	C: Date: 22-Nov-1996 #sequence_revision 22-Nov-1996 #text_change 20-Apr-2000	Qy	411 -----
C: Accession: S71551	C: Function: Cloning and analysis of the DNA polymerase gene from a new hyperthermophil	Db	410 478 TRIDSLAKSFLTRKVKGXKIIITPLFYBIRGATSENPEEEVLKGELAGILLAEGTLRK 537
R: Kakihara, H.; Takagi, M.; Imanaka, T.	A: Description: Submitted to the EMBL Data Library, March 1994	Qy	411 -----
A: Reference number: S71551	A: Molecule type: DNA	Db	411 538 DVEYDSSRKRRISHOYRVEITIGKDEEFRDRITYFIRLFGITPSISERKGTNAVTL 597
A: Cross-references: EMBL:D29671	A: Cross-references: EMBL:D29671	Qy	411 -----
C: Function: <NUCL>	C: Function: as DNA-directed DNA polymerase, catalyzes the Polymerization of DNA at the	Db	411 598 KVAKKVNLYKVEIMDNIESLHAPSVLRGFFEGDGSYVRVRSIVATOGTKNEWKIKLVS 657
A: Note: DNA-directed DNA polymerase KOD	A: Note: DNA endonuclease PI-I, catalyzes the hydrolysis of internal phosphodi	Db	411 658 KLLSQLQIPHQTYTYQYQENGKDRSRYILETGKDGGLFQTLIGFISERKNALLNAIS 717
C: Function: <BN1>	A: Description: as DNA endonuclease PI-I, catalyzes the hydrolysis of internal phosphodi	Qy	411 -----
A: Note: DNA endonuclease PI-PspI	A: Description: as DNA endonuclease PI-II, catalyzes the hydrolysis of internal phosphodi	Qy	411 659 KLLSQLQIPHQTYTYQYQENGKDRSRYILETGKDGGLFQTLIGFISERKNALLNAIS 717
C: Superfamily: DNA-directed DNA polymerase KOD	C: Superfamily: DNA-directed DNA polymerase KOD	Db	411 718 QREMMNLNEGPFYRILSEFNSTETYEGKTYDILTEGTPYYFANGLILHNSLYSIIIIH 777
C: Keywords: endonuclease; hydrolase; nucleotidyltransferase; Protein splicing	C: Keywords: endonuclease; hydrolase; nucleotidyltransferase; Protein splicing	Qy	411 778 VSPDTLNREGKEYDQAPQVHFRKDFQFPIPLGDLLEEROKIKKMKATDPIERK 837
F1: 406-767-851-1388-1670/ Product: DNA-directed DNA Polymerase KOD #status Predicted <MAT>	F1: 406-767-851-1388-1670/ Product: DNA-directed DNA Polymerase KOD #status Predicted <XT1>	Db	411 838 LDYDRAIKILANSILSILPEEWLPLVYEEGEVHFVRLGELDPMRNLKNYRRLYBALVEN 897
F1: 406-766/ Product: DNA endonuclease PI-I (pol) KOD intein 1) #status Predicted <MBT2>	F1: 406-766/ Product: DNA endonuclease PI-I (pol) KOD intein 2 #status Predicted <XT2>	Qy	411 838 LDYDRAIKILANSILSILPEEWLPLVYEEGEVHFVRLGELDPMRNLKNYRRLYBALVEN 897
F1: 767-851/ Domain: DNA-directed DNA Polymerase KOD #status Predicted <XT1>	F1: 767-851/ Domain: DNA-directed DNA Polymerase KOD #status Predicted <XT2>	Db	411 838 LDYDRAIKILANSILSILPEEWLPLVYEEGEVHFVRLGELDPMRNLKNYRRLYBALVEN 897
F1: 1388-1670/ Product: DNA endonuclease PI-I (pol) KOD intein 2) #status Predicted <XT3>	F1: 1388-1670/ Product: DNA-directed DNA Polymerase KOD #status Predicted <XT3>	Qy	411 838 LDYDRAIKILANSILSILPEEWLPLVYEEGEVHFVRLGELDPMRNLKNYRRLYBALVEN 897
F1: 406-767/Cross-link: peptide (Arg-Ser) #status Predicted	F1: 406-767/Cross-link: peptide (Arg-Ser) #status Predicted	Db	411 838 LDYDRAIKILANSILSILPEEWLPLVYEEGEVHFVRLGELDPMRNLKNYRRLYBALVEN 897
Qy	Score: 2764.5; DB: 2; Length: 1670;	Qy	411 838 LDYDRAIKILANSILSILPEEWLPLVYEEGEVHFVRLGELDPMRNLKNYRRLYBALVEN 897
Best Local Similarity: 35.9%; Pred. No. 4.5e-134;	Matches: 599; Conservative 83; Mismatches: 88; Indels: 899; Gaps: 4;	Db	411 838 LDYDRAIKILANSILSILPEEWLPLVYEEGEVHFVRLGELDPMRNLKNYRRLYBALVEN 897
Db	Db	411 838 LDYDRAIKILANSILSILPEEWLPLVYEEGEVHFVRLGELDPMRNLKNYRRLYBALVEN 897	
Qy	1 MIFDTYITRDKGPITI FRKENGKTRIEFLDPHFTPYTALLKDDSAIBIKATKGKRG 60	Db	411 838 LDYDRAIKILANSILSILPEEWLPLVYEEGEVHFVRLGELDPMRNLKNYRRLYBALVEN 897
Db	1 MILDTDYITEDGKPVIRKPKENGKTRIEFLDPHFTPYTALLKDDSAIBIKATKGKRG 60	Qy	411 838 LDYDRAIKILANSILSILPEEWLPLVYEEGEVHFVRLGELDPMRNLKNYRRLYBALVEN 897
Qy	61 KIVRYYDAVYKVKCKFLGRDVEWVKLQFLFEPHPODVAFLRKGRHPIHEDYDIFPARY 120	Db	411 838 LDYDRAIKILANSILSILPEEWLPLVYEEGEVHFVRLGELDPMRNLKNYRRLYBALVEN 897
Db	61 KIVRYYDAVYKVKCKFLGRDVEWVKLQFLFEPHPODVAFLRKGRHPIHEDYDIFPARY 120	Qy	411 838 LDYDRAIKILANSILSILPEEWLPLVYEEGEVHFVRLGELDPMRNLKNYRRLYBALVEN 897
Qy	121 LIDKGAI PMEGDEELKMAFDIETFYEGDEFKGKBEIMISYADEEARYITWNIDLPY 180	Db	411 838 LDYDRAIKILANSILSILPEEWLPLVYEEGEVHFVRLGELDPMRNLKNYRRLYBALVEN 897
Db	121 LIDKGAI PMEGDEELKMAFDIETFYEGDEFKGKBEIMISYADEEARYITWNIDLPY 180	Qy	411 838 LDYDRAIKILANSILSILPEEWLPLVYEEGEVHFVRLGELDPMRNLKNYRRLYBALVEN 897
Qy	181 VDVSNSNPEMILFREVQYRKFLRQPKLDFPQYDPAIRKREHGAVIDYEEFLPFAEY 240	Db	411 838 LDYDRAIKILANSILSILPEEWLPLVYEEGEVHFVRLGELDPMRNLKNYRRLYBALVEN 897
Db	181 VDVSNSNPEMILFREVQYRKFLRQPKLDFPQYDPAIRKREHGAVIDYEEFLPFAEY 240	Qy	411 838 LDYDRAIKILANSILSILPEEWLPLVYEEGEVHFVRLGELDPMRNLKNYRRLYBALVEN 897
Qy	241 PKTHRMGDSFAVEIKGKTHEDLFPIVYRARTINLPTYLAEAVYAVLGTKSKLGAEBIAAI 300	Db	411 838 LDYDRAIKILANSILSILPEEWLPLVYEEGEVHFVRLGELDPMRNLKNYRRLYBALVEN 897
Db	241 PKTHRMGDSFAVEIKGKTHEDLFPIVYRARTINLPTYLAEAVYAVLGTKSKLGAEBIAAI 300	Qy	411 838 LDYDRAIKILANSILSILPEEWLPLVYEEGEVHFVRLGELDPMRNLKNYRRLYBALVEN 897
Qy	301 WETBESMKLQYSMEDARATYELGKFFPMEAEKLIGOSWVDSRSSTGNIYEWYLL 360	Db	411 838 LDYDRAIKILANSILSILPEEWLPLVYEEGEVHFVRLGELDPMRNLKNYRRLYBALVEN 897
Db	301 WETBESMKLQYSMEDARATYELGKFFPMEAEKLIGOSWVDSRSSTGNIYEWYLL 360	Qy	411 838 LDYDRAIKILANSILSILPEEWLPLVYEEGEVHFVRLGELDPMRNLKNYRRLYBALVEN 897
Qy	361 RAYERNLAPKNPDEEYRRRQRTTLYGGVYKEPEGLWENITYLDFRC----- 410	Db	411 838 LDYDRAIKILANSILSILPEEWLPLVYEEGEVHFVRLGELDPMRNLKNYRRLYBALVEN 897

Qy	605 RITTRGLEVRRDWSEIAKETOAULKPAILKEDSVAKEVIEVKVYEEIAKYQVPLEKLY 664	Db	453 FFKRILMLIEKRREKLVELKNLSPESSEYKULDIKQOTLKVLTNSFYGGMGNLLARWYC 512
Db	1498 KITRGLBIVRRDWSEIAKETQARVLEALLKGDVKEAVERVTEBKSLXEVPEKLV 1557	Qy	510 KCEADESVIATWGRYYIEMTIKELEBKFGPKVLYADTDGYATIPEGKPTIKKAKEFKLK 569
Qy	665 IHEQITKVLSEIAKETGAIGKIKYRPGTISIYSTVLRGSGKISDRVILLSEYD 724	Db	513 HPCBAAATIAGRPIRTBAK-IAESMGKVKYQGTDTSFPTVYKAGMTEDVDR--JIDK 567
Db	1558 IHEQITRLLDKYDRTATGPVAVAKRLLAARGVKIRPGTISIYSTVLRGSGRIGDRAFTEDF 1617	Qy	570 INSKPLGILLELEYEGFLRGFFYAKRKYAVIDEDEGRITTRGLEVVRDWSSEIAKETOKV 629
Qy	725 PKCHKYDDYIENQULPAVRLILEAFYRKEKDLYKOSKQVGLDMLK 773	Db	568 LHEBP-TIQIEDEYYSAIFFYKERYAGLTDGRLVKGLEVRGDWCELAKKVOREV 625
Db	1618 PTXHKYDDAEYYTENQVLPAVERILRATPGYRKEKDLYQKTRQVGLSALWK 1666	Qy	630 LEAIKEDSVERAKEIVKDVERIAKTYQVPLKLYTHEQITYDLSYEKAYAIGPHVIAKRL 689
RESULT 8			
A9312 DNA polymerase B1 (polB) homolog - Archaeoglobus fulgidus			
C.Species: Archaeoglobus fulgidus			
C.Date: 05-Dec-1997 #text_change 05-Dec-1997 #text_change 22-Oct-1999			
C.Accession: A69312			
R.Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson, J.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; McDonald, L.			
Nature 390, 364-370, 1997			
A.Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.; Smith, H.O.; Woese, C.R.; Venter, J.C.			
A.Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeon Archaeoglobus fulgidus.			
A.Reference number: A69312			
A.Accession: A69312			
A.Status: Preliminary; nucleic acid sequence not shown; translation not shown			
A.Molecule type: DNA			
A.Residues: 1-781 <KLE>			
A.Cross-references: GB:AE0001070; GB:AE000782; NID:92689393; PIDN:AA80741.1; PID:9265013			
C:Superfamily: herpesvirus DNA-directed DNA Polymerase			
Query Match Score 1301; DB 2; Length 781;			
Best Local Similarity 37.4%; Pred. No. 2.3e-59; Matches 205; Conservative 155; Mismatches 266; Indels 90; Gaps 21;			
Qy	2 IFPDTDXTIDGKPIIRFKKENGEFKTLEDPHQFQPYIYALKDDSAIDEKAIGKERHKG 61	Db	62 IYVDAVAKVKKFKLGRDVEWKLIFHPDQDYLKQVYDPPVYDQFVAYDNDEPQVYD 121
Qy	8 LIDADYETIGKAVVRWMCQDQGIVAYDNDEPQVYDQFVAYDNDEPQVYDQFVAYDN 64	Db	65 VIKLKSFFKAQKTLGRVEGTYIYAHPOHVKLDRYLSQFG---DVERADIPAYRVL 121
Qy	122 IDKGL---IPMEGDE-----ELKLMAPDDEIYHEG-DEFG 153	Db	122 IDEDLACNDGTAIEGEKQGGVIRSYKIEKVERIIPMEPEKLMYEDCEMILSFCMPPE 181
Qy	122 IDEDLACNDGTAIEGEKQGGVIRSYKIEKVERIIPMEPEKLMYEDCEMILSFCMPPE 181	Db	154 KGBIIMSYADEEARVFTWNIDLPVYDVSNEREMIKRFVQIVREKDPVLYTNGDN 213
Qy	182 KDPPIIVSVKTNDDDEI-----LGDERKIKISDFEVKLIKSDFVLYTNGDN 230	Db	182 KDPPIIVSVKTNDDDEI-----LGDERKIKISDFEVKLIKSDFVLYTNGDN 230
Qy	214 FDPYLLKRAEKGTVLGRDKHEPEPKIHRMGDSFAVEIKGRTHFDLFPVVRTINLP 273	Db	214 FDPYLLKRAEKGTVLGRDKHEPEPKIHRMGDSFAVEIKGRTHFDLFPVVRTINLP 273
Qy	231 FDPYLLKRAEWNIPDVGIGSN---VVFRGG---RPKTGRANVDIAMIKRISDK 284	Db	231 FDPYLLKRAEWNIPDVGIGSN---VVFRGG---RPKTGRANVDIAMIKRISDK 284
Qy	274 TYLEAVTEAVLGKTK-SKLGAEEIAAIWETEBSMVKLAQYSMEDARAYELGKEFPE 332	Db	274 TYLEAVTEAVLGKTK-SKLGAEEIAAIWETEBSMVKLAQYSMEDARAYELGKEFPE 332
Db	285 IKKLENTEAFLGKTKIADEIAKDIYWSGEK-EKVLNTARQDANTYLIAKELLPH 343	Db	285 IKKLENTEAFLGKTKIADEIAKDIYWSGEK-EKVLNTARQDANTYLIAKELLPH 343
Qy	333 AELAKLIGOSYNDVSRSSTGNLVEWYLVRAYERNELAPNKPDEEYRRRLRTTYLGGY 392	Db	333 AELAKLIGOSYNDVSRSSTGNLVEWYLVRAYERNELAPNKPDEEYRRRLRTTYLGGY 392
Db	344 YELSKMIRLPVDDVTRGRKQDWMILSEAKIGIAPNPEHAE-----SYEGAFV 396	Qy	393 KEPERGLKENTYLDPROLYPSITVHNNSPDTLREGCKN-YDVAITYGKPKDFPG 450
Qy	397 LEERGIGHENACDFASMYPSIMAFNPSDTCY-APEVGHKFKSPDG 452	Db	397 LEERGIGHENACDFASMYPSIMAFNPSDTCY-APEVGHKFKSPDG 452
Qy	451 FIPSILGELITNRQEIKKRMK-ATIDPIEKNMDFQRAVKLHANSYYGNGYPKARWYS 509	Db	451 FIPSILGELITNRQEIKKRMK-ATIDPIEKNMDFQRAVKLHANSYYGNGYPKARWYS 509
Query Match Score 1291; DB 2; Length 1634;			
Best Local Similarity 22.7%; Pred. No. 2e-58; Matches 159; Conservative 159; Mismatches 374; Gaps 23;			
Qy	2 IFPDTDXTIDGKPIIRFKKENGEFKTLEDPHQFQPYIYALKDDSAIDEK-AIK 55	Db	13 LIINTYKTEDKAVIYLINS---ILKDRDFKPYVYELHEKEVYENEDIEKIKDFLLK 68
Qy	56 GBRHKIVRUVDAVKVKK---FLGRDVWVKLTFEPHDVPAIRGKREHPAVDIVEYDI 114	Db	56 GBRHKIVRUVDAVKVKK---FLGRDVWVKLTFEPHDVPAIRGKREHPAVDIVEYDI 114
Qy	69 ND---LKEKEVNEYKTKILRKERVKIKIATHPKVPLR-KIKECEVKEIYHDI 123	Db	69 ND---LKEKEVNEYKTKILRKERVKIKIATHPKVPLR-KIKECEVKEIYHDI 123

Qy	115 PFAKRYLIDKGILPM-----EGDE-----ELKLMADFIEDTFYHEGDEFGKGE-TIMIS 161	494	-----	493
Db	124 PFAKRYLIDNEIIPNTWYDFFENKPKVS-BIPKUKSVAFDMEVNRDTEPNPBDPLMAS 183	494	-----	493
Qy	162 YADEESEARVITIWKNIDLIPYDVSNERMIKRTQIVVERKDPLVITYNGDNFDPLIK 221	494	-----	493
Db	184 FWDENGKIVITYGEFNPNIEWVNKEKEIJKLIELEY- -DVIYTNGDNFDPLKA 241	494	-----	493
Qy	222 RAEKLGVTLLIGRDKHEPRPKHRMGDSFAVEIKGRHFDLFFPVYVRRATINIPTTYLEAVY 281	494	-----	493
Db	242 RAKYGIIDNLGKGE- -ELKIKRGGMMEYRSYIPIGRVIDLIPYISRRLKLTKTYLDDVV 299	494	-----	493
Qy	282 BAVLGKTKSKLGAEBIAIAWETESEMKLQAQYMSMEDABATYELGKEPPMEEAFLAKLGQ 341	494	-----	493
Db	300 YNLFGTEKLEKMKIPIHTRKIVDWNND- -KTLIEYESQDACYTKYQKGYFPFMLEMSRIVNQ 357	494	-----	493
Qy	342 SVWDYRSSTGNLNEWYLLERVAYEENELAPNKPDEEEFRRRLTETLLOGTYKEPERGLWE 401	494	-----	493
Db	358 TPEELTRMSSQMVYLLMKRAFKNMIVENPKDPEEEVRRVLTTEYEGCVYKEPEKMF 417	494	-----	493
Qy	402 NITYLDFRC-----	494	-----	493
Db	418 DIISMPFRCHPKGTIVKVKYKGVNIEDYKEREGNYL/GIDGQWQKVKYKQYEFGELNNV 477	494	-----	493
Qy	411 -----	494	-----	493
Db	478 GLKCTPNHK1PLRYKTKKKTNKNDYLVRDIYAKSLLTKEFGESKLICKDFETIGNYK 537	494	-----	493
Qy	411 -----	494	-----	493
Db	538 YINMD3DFILKSELIGLLAEGHLLRDTBYFDSRGRKKRISHQYRVEITVNEDEDKDFI 597	494	-----	493
Qy	411 -----	494	-----	493
Db	598 EKIKYIFKLNELYVRRKGKTRATLGZAKDITYKLEBILKRNKEYTKVLLPNAILRGFEE 657	494	-----	493
Qy	658 GDGYNTVRAVVNNQGNNYDKIKFIAASLLDRGKLYKSYTYSYBERGRKLYKRYVIEF 717	494	-----	493
Db	718 SKGLIKFSLISFISRKNNNLNEITRQKTYKGDYFYDLDVVSLESYKGEXVYDL 777	494	-----	493
Qy	778 TLEGRPYFYFANGLTHNSLYFSIIISNISPDTLCECCK--DYESEKLGHWFKCKEGL 835	494	-----	493
Db	842 IPSILGELITMRQEIKRKM--ATIDPI--EKKMLDYRORAVKLA- -	494	-----	493
Qy	836 IPEKTLRNLIERNIKERMKMAETIGINEEYNNLDYEQRSKILANSILPDEYLTIEE 895	494	-----	493
Qy	896 DGIKVKVTKGEYIIDDLMRKHKDKIKKFGSISILETBNKLTFSKDITKCKEIKKVKALIRH 955	494	-----	493
Qy	956 PYFGKAYKIKLRSGRTIKVTRGHSLFVYENGKIVEVYKDDVYFGLDILVVKLTCYDKEV 1015	494	-----	493
Db	1016 VNIPIKRJINADEEIKLIVLTKDIAFFVVLKKTLEDIENNKLUFDDCILYKLGL 1075	494	-----	493
Qy	1076 LIDNIKIKKINKYDIKLDEEKPKAYKCYFDITVIEHGNFKKGRNCIYIKIDYANIPD 1135	494	-----	493
Qy	1136 KEFEDCEIGAYSGKINALLKLDEKLAKLKEGLFFVTRGRUKKQKLUKGETVYEISVYKSPEY 1195	494	-----	493
Db	1171 FVYRFLIDSVIPBELEFGQVENDSAPSVTTDVTVEVTRVQSTGAGHLDIFLDI 176	494	-----	493
Qy	1343 ETFHSEG-DEFGKGEIIMISYA-DEEARVITMKNIDLPYDVVSNEREMIKRFVQFVRE 200	494	-----	493
Db	177 EVRNPHGMPDPEKEDEIMVGANGNGYESVISTAGDHLDFVEVEDERELLERFAEIVD 236	494	-----	493

Qy 1196 QKEIABTFKEVFGAGSMYKDKVTMDNKIVLVLKX1FKCGDOKKHIPEELFLASESYIK 1255

Db 1256 SFLDGFLLAKKNSHRGTSTMAKDEKYLQNLMLFLNLYGIFTPTVKNGKGKYLTLNPKY 1315

Qy 1316 GTVKDMLDEVKEIAFEYSGYYDLSVEDNENFLVNTYAHNSVYGTIAFPARFYSRE 1375

Db 1320 CAESVIAWGBHYIETNTIKEEKGKFLYADTGFYATIGEK- -PETIKKAKEPELKY 569

Qy 1376 CAEITYLGRKYIETVKE AEKCFKVLXIDTGFYA-IWKENKSKEBLIKAMEFVEY 1433

Qy 570 INSKPLGLREBEGRYLGFVAKKRYAVIIDEGRITRTRGLEVVRDMSIEAKTOAKV 629

Db 1434 INSKPLGTMELLEFGYFKRGIFVTKYKRYALIDENGRTVKGLEFVRDMSNIAKTQRV 1493

Qy 630 LEAIIKEDSVKAELVQDVVEEIAKYQVPLEKLVTHEQITKDLSYEKAIGPHVIAIKRL 689

Db 1494 LEALVVEGSTEKAKLQIYDVKDLEKKTQDLYTQLTKDPEKEYTAPHYIAKRL 1553

Qy 690 AAKGKIVKRPGLTISIVLRSGKIDRVTILSEYDPKGKHYDPEYDQVLPALVRLINE 749

Db 1554 NREGRRIKVGDLLGIVLKVGKTSISERAKLPEEVD- -IDDDVNTYIDNQILPPVLRINE 1611

Qy 750 AFGYRKEDLKYQSOXKQVGLDAWK 773

Db 1612 AVGSKNELKEGA-OLTLDRFK 1634

RESULTS 10

C69028 DNA-dependent DNA polymerase family B (PolB1) - Methanobacterium thermoautotrophicum : DNA-dependent DNA polymerase family B (PolB1) - Methanobacterium thermoautotrophicum : C:Species: Methanobacterium thermoautotrophicum

C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999

C:Accession: C69028 R:Smith, D.R.; Boucette-Stramm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldridge, T.; Qiu, D.; Spadato, R.; Vicaire, R.; Wang, Y.; Wierbowski, J.; Gibson, R.; Jiwani, T.; Qiu, D.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.; J. Bacteriol. 179, 7135-7155, 1997

A:Title: Complete genome sequence of *Methanobacterium thermoautotrophicum* Delta H: fun

A:Accession number: A69000; PMID: 9803754; PMID: 931463

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Cross-references: GB:AE000088; GB:AE000666; NID:92622304; PIDN:AB05697.1; PID:92622

A:Experimental source: strain Delta H

C:Genetics:

A:Gene: MTH1208

Query Match 30.0% Score 1209.5; DB 2; Length 586;

Best Local Similarity 42.9% Pred. No. 7.7e-55; Indels 57; Gaps 10; Matches 255; Conservative 107; Mismatches 176;

Qy 1 MIFPTDYITKDGKPIIRIFKKB --NGFKFIEUDPHQPYIYALLKD-DSAIDEIKAKIG 56

Db 7 VLLDIDYVTDENVIRLEFGDKDGSNEPIIAHRSRSPYIAKPTDOLCECLREBEELEL 66

Qy 57 ERHSGKIVRVTDAVYKK- KFLGRDVEYKLLIFEPHQDYLALRKGIRBPAVIDYEDIP 115

Db 67 BK- - - - -LEVKEMDLGRPTEVTRIIFRHPQDVPKIRDLESVRDIRESDIP 116

Qy 116 FAKRYLIDKGЛИME-----GDEELKLMAFDI 142

Db 1171 FVYRFLIDSVIPBELEFGQVENDSAPSVTTDVTVEVTRVQSTGAGHLDIFLDI 176

Qy 1433 ETFHSEG-DEFGKGEIIMISYA-DEEARVITMKNIDLPYDVVSNEREMIKRFVQFVRE 200

Db 177 EVRNPHGMPDPEKEDEIMVGANGNGYESVISTAGDHLDFVEVEDERELLERFAEIVD 236

RESULT 11
 JC7382 DNA-directed DNA polymerase (EC 2.7.7.7) B3 - Sulfurisphaera ohwakuenensis
 C;Species: Sulfurisphaera ohwakuenensis
 C;Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 31-Dec-2000
 C;Accession: JC7382
 R;Iwai, T.; Kurokawa, N.; Itoh, Y.H.; Kimura, N.; Horiuchi, T.
 DNA Res. 7, 243-251, 2000
 A;Title: Sequence analysis of three family B DNA polymerases from the thermoacidophilic
 A;Accession: JC7382
 A;Molecule type: DNA
 A;Residues: 1-781 <DNA>
 A;Cross-references: DDBJ:AB032376
 C;Comment: This enzyme has both 3'-5' exonuclease and polymerase activities, and plays a
 C;Genetics:
 A;Gene: B3
 C;Keywords: DNA replication; exonuclease; metal binding; nucleotidyltransferase

Query Match 29.9%; Score 1204.5; DB 2; Length 781;
 Best Local Similarity 36.6%; Pred. No. 2e-54;
 Matches 294; Conservative 141; Mismatches 282; Indels 87; Gaps 21;

Db 2 IFDDIXTGDGPPIRIRK-KENGEFKEIQLDPHQFQPYIYALIKDSDAIDEK--AIGER 58
 Db 7 ILDFSYDYYENKPYIYIWIDKEGNRYVILEKFRPYALYVDSYNTDEIRKEILKLSK 66
 Qy 59 HGKIVRVVDAVKYKEKFGDGRDVWKLJLFEHPDVPALRGKLRHPAVIDYEYDPEAK 118
 Db 67 PYSPITSDV--SEKKYFSPSPVYKLTETVIFAYVYRDEVAKIKSYVLEADTRFM 124
 Qy 119 RYLDKGJLPM---ESEDE-----ELKUMAFDIEFYH 147
 Db 125 RYSDINKLPFYWIEAEVIEKENNFRKVKYELKINKLYEDKIPLKVLADIEVY-- 182
 Qy 148 EGDFEG----KGEIIMISYADEEARYITWKNDLPIYDVYNSNEREMIKRFQVIREKD 202
 Db 183 -NKYGSPPNPRDPVIIIGWVTKFEGKOFGLADKYD-----DIRAIREBINFQTYD 231
 Qy 203 PDVLIITYGDNFOLPILYI-KRAEKLGLTLLGRDKEHBPKIHMRMGDSFAVBIKGRHEDL 262
 Db 232 PDIIVGNTNNFDPYIILBRANTRGIRDVGR-RVNEBPS--QGYGHYSITGRANVDL 287
 Qy 263 FPVYRRTINLPYTL--AVYBAVIGKTKSUGA-EIIAAI-WETESSMKKLAQYSMEDAR 319
 Db 288 YGFAQSIIDEVKRTLENNADYLGLVLPKERRQTERTIVEWDPYKWDDEKKRDLKYNDDAK 347

RESULT 12
 B56277 DNA-directed DNA polymerase (EC 2.7.7.7) II - Pyrodictium occultum
 C;Species: Pyrodictium occultum
 C;Date: 03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change 20-Jun-2000
 C;Accession: B56277
 R;Uemori, T.; Ishiro, Y.; Doi, H.; Kato, I.
 J. Bacteriol. 177, 2164-2177, 1995
 A;Title: The hyperthermophilic archaeon Pyrodictium occultum has two alpha-like DNA pol
 A;Reference number: A56277; MURID:9523290; PMID:7721707
 A;Accession: B56277
 A;Molecule type: DNA
 A;Species: Pyrodictium occultum
 A;Residues: 1-803 <UEM>
 A;Cross-references: GB:D38574; NID:9807829; PID:BA07580.1; PID:9807830
 C;Superfamily: herpesvirus DNA-directed DNA polymerase
 C;Keywords: nucleotidyltransferase

Query Match 29.6%; Score 1190; DB 2; Length 803;
 Best Local Similarity 33.7%; Pred. No. 1.2e-53;
 Matches 281; Conservative 164; Mismatches 278; Indels 110; Gaps 23;

Db 1 MIFDDDY-----ITKDGKPIRIFKENGFEKIEIDPHFPPYIALLKD-D 45
 Db 8 VLLDSYYEILGKEPVILGWTIDGKRVYL-----LDHREFPPYPAVIALARGYE 55
 Qy 46 SAIDBEKA1KGERHKGIVR--VYDAVKYRKKEFGLDRDVWKLJLFEHPDVPALRGKIRKH 103
 Db 56 DMVEIIIA--SIRPLSVVSKPSPIDAKPLDKYFGRPAKAKTTMIPESVRAVYKII 113
 Qy 104 PAVIDIYEDIPPAKRYLIDKGH-----IPME-----
 Db 114 EGVEDSLBADIIFRAMYLLDGRLYPFTYRIPBDAENGPGRVDRYVKVAGDPEPLADI 173
 Qy 135 -----LKLMADFJETFYEGD-EFGKGEITMISYADEE-FARVLTWNKNULPYDVVS 185
 Db 174 TRIDLPPMELVAFIEVYSSRGSPNPAIDPVTIVSLRSEGERLIAEGHD----- 225
 Qy 186 NEREMIKRKFQVIVREKDPDVILITYNGDNFDLPIYLIKRAEKLGLTLLGRDKEHPEPKIHR 245

Qy 15 IIRIFKENGEPFKIELDPH-FOPYIYALLKODSAIDEI----- KAIKGERHGIK 61
 Db 88 IIRIFGTTKEGHSVCCNTHGPFYFTICPPGMGPDDISHPHQTLERGRMEANRNSNVGK 147
 Qy 62 IRRVVDAYKVK-----KKELGRDVEWKLIFEPDQVPAALRGKIREHPAVID--- 108
 Db 148 FVRRIEWMQRREIIMYQNSDPEFL-----KIVVALPTMASCRG-TLDRGIGLDLG 199
 Qy 109 -----IYEDIPPAKRYLIDKGLI----- 127
 Db 200 MKSFLTYESEVNLALRFMIDCNVGENWIGIPAGKYYKATKSLSYCQLEFDCLYSELISH 259
 Qy 128 PNEGD---EEKLIMAPDIEFTYHEG-DEFGKGETLIMISADEEPAVITWNKIDLPY- 180
 Db 260 APEGEYSKMAPRILSFDIEAGRKHFFPEPTHDLPVQI-----ANVYLOQEDQPI 312
 Qy 181 -----VDVV--SNEREMIKRFVQIVREKDPDVLYTNGDNFDLPLVLIKRAE 224
 Db 313 RNMVTLKSCSPVIGDVMPPFETEYLLAARDIFIREVDPDIIIGNICPDLPILLER 372
 Qy 225 KLGVT--LLGGDKEHP-----EPKIHRMGDSFAVEIKGRIHFDLFPPVVRERINTLP 274
 Db 373 NUKIAEFPILGRIRNSRVRVXRDTFSSRQYGTRESKEVAVEGRVTFLLOVMDYKUSS 432
 Qy 275 YLEAVTEAVIGKTKSKLGAEBIAAWE-IEBESMKLQASMSMEDARATEVLGKE- FPFM 331
 Db 433 YSLNSVSSHFUSEQEDVHHSITSDIQNNGAETRERLAVYCLDKAYLPORLDDLMFTYN 492
 Qy 332 EABLAKLIGOSV-WDVSRSSTGNLYEWYLRLRVAERNELAPN-KPDEEYRRRLRTYLG 389
 Db 493 YVEMARYTGVPSFLISRGQIKVL-QLIRRARQNLWPNKAOGSE----QGTFPG 546
 Qy 390 GYKKEPERGLMEN-ITYLDFERCLYPSITIVTHANVSPDTLE REGCKNYDNAF----IYG 441
 Db 547 ATVLEARAGFYEKPIATLDFAASLYPMIMAYNLCYCTLVIPEDARKLNLPPESYNRTPEG 606
 Qy 442 YKPCKD--PPFPIPLIGELTMRDPIKPKATIDPIERKMLDYQRAVKLHANSYTG 499
 Db 607 BTFVKSNLQKGTLPEELLTARTKAKADIKEADPILEKAVDGRQLALKISANSVYGP 666
 Qy 500 NGYPKARWYKSEKAEVSHTWYRHELTMEKIEEPF---GF---KVLYADTDGFYATI 551
 Db 667 TGTATGOLPCLISSVTSISQRMEMHTKLIVEEFTTINGYERAEVYFGDTSVWTOF 726
 Qy 552 PGEKPETIKKEAFKUYTINSKPGHLELEYEGFTLRFQFVAKRYAVI-----DEEGRI 606
 Db 727 GVSAYEAMNLNGREALEHISGFTTPKLEFEKVTYPLLISKRYAGFWTREDNEDIM 786
 Qy 607 TTRGLEYVRDWESETAKETOAKVTEALKEDSVKEAVIYKDVYELAKYQVBLEKLVH 666
 Db 787 DTKGIETRVRDNCLLYRNLYNDCLKILIDPDAVQVYKNAISDLINNRMDSLILYT 846
 Qy 667 EQUITKOLSEYKAIGAIPHVAIAKLAAGKIGKVRP-GTIIISIVLRS---GKISDRVY 718
 Db 847 KGLTRKTDYEVKAHVELAERMRKDATAAPNVYDPPVITIAAKGKAYAVERSEDPY 906
 Qy 719 LSEYDPKXKHYPDYYIENQVLPAVLRLIE 749
 Db 907 VLENNTP---IDPHYLQNIQSKPLTRLIFE 933

RESULT 15
 S75407 probable DNA-directed DNA polymerase (EC 2.7.7.7) - Sulfolobus solfataricus
 N; Alternative names: protein c01041
 C; Species: Sulfolobus solfataricus
 C; Date: 11-Oct-1997 #sequence_revision 24-Oct-1997 #text_change 20-Jun-2000
 C; Accession: S75407
 R; Sensors, C.W.; Klenk, H.P.; Singh, R.K.; Allard, G.; Chan, C.C.Y.; Liu, Q.Y.; Penny, S.
 M; Microbiol, 22, 175-191, 1996
 A; Title: Organizational characteristics and information content of an archaeal genome: 1
 A; Reference number: S73076; MUID:97055432; PMID:8899719

Qy 16-88 Score: 676.5 ; DB 2 ; Length: 764 ;
 Best Local Similarity: 25.33 ; Pred. No. 2.6e-27 ;
 Matches: 206 ; Conservative: 168 ; Mismatches: 312 ; Index: 127 ; Gaps: 24 ;
 Qy 2 1FDTDYTPKDGKPIIRIFK-KENGEFKIELDBHFQPYIYALLK-DDSAIDEIKAKGGER 58
 Db 7 1LDSEYFKGNTPLVYIWSVSDDEGNSSVIYNSNFRPYFYIYEGNNNEIINIK----- 60
 Qy 59 HKIVRIVUDAVKTKKELGRDPEWKLIFPQDVALRGKIREHAPVIDIYEYD----- 114
 Db 61 -KNCEAQIQTWKYKRYLGNITDALLIQTSTPQIKCREKISELNNKIGLFDADRYTM 118
 Qy 115 -----PFA-----KRYLIDKGLIPMGEQ-BEFLKUMAFDIBTFYH 147
 Db 119 RYSLDEFDLPRTMPEAEVNEYKEDGPKTQYILDILISHLVYEGNMPELRTIGDVFYI----- 176
 Qy 148 EGDFG-----KGBIMTMSYADEEAVRITYWKNIDLPYDVSNEREMIKRFVQITREKD 202
 Db 177 -SKYGSINPRKDPIVMSLNSKEGMQFS-----LDGEGDJKIIRFVDYTLNVD 226
 Qy 203 PDVLTITYGDNFDLPLVYIJKRAEFLGTYLLGDRKEHPEPKTHRMGDSFAV-----EIKG 256
 Db 227 PDIIIVYDSLLEPKWYKITERASSLGKIDIG-----KIGBEVSCTYGHMSISG 276
 Qy 257 RTHFDLFFYVVRTTINLPTTYLEAVYAVLGLKTKSKLGAE--EIAAIWETEBSMKLQAYS 314
 Db 277 RNUVDLGLTLLVNEBSLGHVDLIDV-SNYLGSPSRYSKRYEISRWNDNEQNRRIREYS 335
 Qy 315 MEDARATYELGKEFFPMEBAELAKLIGOSWYDVSRSSTGNLYEWYLRLRVAERNEL2LPNKP 374
 Db 336 IENARASYLLGNYLLSITYSELVKGIVPLDKLSVSWGNRLETSLIRTAKSGLPPIRM 395
 Qy 375 DEEYRRRLRRTYLGKVVKPERGLMENITYLDFRLYPLSPITVHVSPTDLERBCKNY 434
 Db 396 DNPNRSPKIKK---NIIIQPKVGIYDITDYYVWIDISSVYSLVIRKPNIAPDTLVKECDC 451
 Qy 435 DVAPIVYKFCKDFFGTPSTIGELTMQRBIKKMKATDPIERKMLDYRQRAVYLHAN 494
 Db 452 YSSPISNYKFRPSPGILYKTFDELDNSRDSNKVY-----IEELI-----S 493
 Qy 495 SYGYGMPYKARWYKSEKAEVSHTANGRHYETMTEKIEKEKGFKVKYLADTDGFYATIPGE 554
 Db 494 SFNDYHWNARWYSEHIASEDEFENNEIRPFLIDLIKSS-GLDVYLANDLILIFTV----- 548
 Qy 555 KETIKKCKAKEPLKYNKSLPGLLELEYEGFTLRFQFVAKRYAVIYDEEGRITTRGLEVY 614
 Db 549 -GGSRKVNELITKNSLYN-LDYVVKLYFKSLLIVDNNRYAGISSEGPKI-----DIA 599
 Qy 615 RR-----DWSEIETAKETOAKVTEALKEDSVKEAVIYKDVYELAKYQVBLEKLVH 669
 Db 600 RKGEDMNCLCEARNRKRKIBELLSKDVKAIVKSTVILKRGFDFNEELITWAKI 659

Qy 670 TKDLSEYKAIGPHVAIAKLAAGKIGKVRPGLIISVYVLRSSGKISDRV--ILSEYDPK 726
 Db 660 ERDLNEYNNQLPFTVIAKQSGVILSKDSKIGTVKGPLNDRAEPEFLVKE----- 715
 Qy 727 KHKYDDYYIENQVLPAVLRLIEAFGKREKDLK 759
 Db 716 KNRDIIEYYD-QIFFRETKLKPLGVNEESLK 747

Search completed: February 9, 2004, 11:53:19
 Job time: 61 sec

Result No.	Score	Query	Match	Length	DB	ID	Description
1	3446.5	85.6	1829	1	DPOL_THEST		O33845 thermococcus
2	3235.6	83.4	1702	1	DPOL_THELI		P30317 thermococcus
3	3253.5	80.8	773	1	DPOL_THESG		Q56369 thermococcus
4	3224.5	79.6	775	1	DPOL_THES9		P77916 pyrococcus
5	3164.5	78.6	771	1	DPOL_PYRAB		P80061 pyrococcus
6	3131.5	78.3	775	1	DPOL_PYRFU		Q51334 pyrococcus
7	2924.5	72.9	1312	1	DPOL_PRSID		Q59610 pyrococcus
8	2886	71.7	1235	1	DPOL_PYRHO		P74918 thermococcus
9	2809	69.8	1523	1	DPOL_PYRHO		P77933 pyrococcus
10	2782	69.1	1671	1	DPOL_PYRKO		Q9hb84 thermococcus
11	2769.5	68.8	1699	1	DPOL_THEGB		Q9hb05 thermococcus
12	2335.5	63.0	1668	1	DPOL_THEHY		P52025 methanococcus
13	1392	34.6	824	1	DPOL_NETVO		O29753 archaeoglob
14		32.3	781	1	DPOL_ARCFU		Q58235 methanococcus
15	1292	32.1	1634	1	DPOL_METJA		O27276 methanobacter
16	1209.5	30.0	586	1	DPOL_METH		P45346 aeropyrum
17	1138	28.3	784	1	DPOL_AERPE		P09436 oryza sativa
18	697.5	17.3	1105	1	DPOL_ORYSA		P30116 schizosaccharomyces
19	693.5	17.2	1088	1	DPOL_SOYBN		P54158 drosophilae
20	621.9	17.2	763	1	DPOL_STLSD		P15436 saccharomyces
21	676.5	16.8	764	1	DPOL_SULSO		P54747 rattus norvegicus
22	676.5	16.8	1081	1	DPOL_ARATH		P05607 sulfurovum
23	657.5	16.3	1038	1	DPOL_CANAL		P26581 sulfobolus
24	657.5	16.3	1086	1	DPOL_SCBPO		P52331 mus musculus
25	624.5	15.5	1092	1	DPOL_DROME		P97783 mesocricetus
26	621.5	15.4	1097	1	DPOL_YEAST		P30105 plasmidium
27	619.5	15.4	1103	1	DPOL_RAT		P28340 homo sapiens
28	616.5	15.3	872	1	DPOL_SULCH		DR SMART: SM00306; HintNC: 3.
29	616.5	15.3	882	1	DPOL_SULSO		DR SMART: SM00486; HintNC: 1.
30	616.5	15.3	1105	1	DPOL_MOUSE		DR TIGRFAMS: TIGR01443; HintNC: 3.
31	614.5	15.3	1088	1	DPOL_NEMES		DR TIGRFAMS: TIGR01445; HintNC: 3.
32	614.5	15.3	1094	1	DPOL_PLAKF		DR TIGRFAMS: TIGR00592; HintNC: 2.
33	611	15.2	1107	1	DPOL_HUMAN		DR PROSITE: PS00116; DNA_POLYMERASE_B; FALSE_NEG.

ALIGNMENTS

RESULT 1
DPOL_THEST STANDARD; PRT; 1829 AA.ID DPOL_THEST
AC Q33845;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE DNA polymerase (EC 2.7.7.7).
GN P01.
OS Thermococcus sp. (strain TY).
OC Archaea; Euryarchaeota; Thermococci; Thermococcaceae; Thermococcus.
NCBI TaxID=110163;RN [1]
RN SEQUENCE FROM N.A.
RN MEDLINE=98094267; PubMed=9434178;
RN Niehaus F.; Frey B.; Antzianian G.;
RT "Cloning and characterisation of a thermostable alpha-DNA polymerase
RT from the hyperthermophilic archaeon Thermococcus sp. TY.";
RT from the hyperthermophilic archaeon Thermococcus sp. TY.";
RL 204:152-158(1997).
CC -|- CATALYTIC ACTIVITY: N deoxyribonucleoside triphosphate = N diphosphate
CC + (DNA)N).
CC -|- PIM: THIS PROTEIN UNDERGOES A PROTEIN SELF SPLICING THAT INVOLVES
CC A POST-TRANSLATIONAL EXCISION OF THE THREE INTERVENTING REGION
CC (INTEINS) FOLLOWED BY PEPTIDE LIGATION.
CC -|- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-B FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC EMBL; Y13030; CPA73475.1; -.
DR InterPro; IPR0006172; DNA_Pol_B.
DR InterPro; IPR0006172; DNA_Pol_B-dom.
DR InterPro; IPR0006133; DNA_Pol_B_exo.
DR InterPro; IPR003586; Hedgehog_HinEC.
DR InterPro; IPR003587; Hedgehog_HinEN.
DR InterPro; IPR006141; Interin.
DR InterPro; IPR006142; Interin.
DR InterPro; IPR006142; Interin_endonuc.
DR InterPro; IPR000578; Pol12.
DR Pfam; PF00136; DNA_Pol_B.
DR Pfam; PF03104; DNA_Pol_B_exo; 1.
DR PRINTS; PRO0379; InterIN_.
DR SMART; SM00305; HintNC: 3.
DR SMART; SM00486; HintNC: 1.
DR TIGRFAMS; TIGR01443; HintNC: 3.
DR TIGRFAMS; TIGR01445; HintNC: 3.
DR TIGRFAMS: TIGR00592; HintNC: 2.
DR PROSITE; PS00116; DNA_POLYMERASE_B; FALSE_NEG.

DR	PROSITE; PS50818; INTEIN_C_TER; 3.	Db	781 VSPDTLREGCKNYDVAPIVGYKFFCKDFFGFFPSPSTLGELITMRQEIKKMKATIDPIEKK 840
DR	PROSITE; PS50819; INTEIN_ENDONUCLEASE; 2.	Qy	481 MLDYRORAVLHA- 493
KW	Transferase; DNA-directed DNA Polymerase; DNA replication;	Db	841 MLDYRORAVLHA- 493
KW	DNA-Binding; Hydrolase; Endonuclease; Autocatalytic cleavage;	Qy	494 -----
KW	Protein Splicing.	Db	901 VDNIAFSLNKESKSKBIIKVKALLRHKYKGEAYEVELNSGRKIHITRGSLSFTNRNGK 960
FT	CHAIN 1 409 DNA POLYMERASE, 1ST PART (POTENTIAL).	Qy	494 -----
FT	CHAIN 410 769 INTEIN I.	Db	961 KBIWIGEEVKVQDILIVPKVVKLINEKEAVINIPELISKLPDDETDADVWMTPVKGKGNFFK 1020
FT	CHAIN 770 855 DNA POLYMERASE, 2ND PART (POTENTIAL).	Qy	493
FT	CHAIN 856 1392 INTEIN II.	Db	961 KBIWIGEEVKVQDILIVPKVVKLINEKEAVINIPELISKLPDDETDADVWMTPVKGKGNFFK 1020
FT	CHAIN 1393 1441 DNA POLYMERASE, 3RD PART (POTENTIAL).	Qy	493
FT	CHAIN 1442 1598 INTEIN III.	Db	961 KBIWIGEEVKVQDILIVPKVVKLINEKEAVINIPELISKLPDDETDADVWMTPVKGKGNFFK 1020
FT	CHAIN 1599 1829 DNA POLYMERASE, 4TH PART (POTENTIAL).	Qy	494 -----
SQ	SEQUENCE 1829 AA; A113ABBC57BB9CB3 CRC64;	Qy	494 -----
Query Match	85.6%; Score 3446.5; DB 1; Length 1829;	Db	1021 GMRLTJKWIFGEESKRITPRYLFLLEEGFVKLPLRGYEVTDWEGLKEYRQLYVEKLVK 1080
Best Local Similarity	42.2%; Pred. No. 1.2e-179;	Db	1021 GMRLTJKWIFGEESKRITPRYLFLLEEGFVKLPLRGYEVTDWEGLKEYRQLYVEKLVK 1080
Matches	771; Conservative 0; Mismatches 3; Indels 1055; Gaps 3;	Qy	494 -----
Qy	1 MIFDTDYITKDGKPIIIRIPIKKEGEPEKIELDPLPHFOPYIYALLKODSADEIKAKGERHG 60	Db	1081 NLRYNGNKRKREVLVRFNDIKDSVSCPRKELEEWKIGTXKGFRXXKICILKVYDFFGFLGYY 1140
Db	1 MILDOTDITKDGKPIIIRIPIKKEGEPEKIELDPLPHFQYIYALLKODSADEIKAKGERHG 60	Qy	494 -----
Qy	61 KIVRVRVDAVKYKVKKEFGRDVEVKLJFEHPDVPALRKIKEHPDVPALRKIKEHPDVPALRKI 120	Db	1141 VSEGYYGAQKNTGCMSSVTKLYNNENPNVLKDMKNAIAEKFFGKVVGKNCVTDIPKRMAYL 1200
Db	61 KIVRVRVDAVKYKVKKEFGRDVEVKLJFEHPDVPALRKIKEHPDVPALRKIKEHPDVPALRKI 120	Qy	494 -----
Qy	121 LIDKGLIPMEGDBELKLMADIEETPYHGSDFGKGBITMISYADEBEEAIVTWNIDLPY 180	Db	1201 LAKSLCGVTAENKRIPIIIFDSSEPYRWAFLRAYFVGDGIDHPSKRLSLSTKSELLANOL 1260
Db	121 LIDKGLIPMEGDBELKLMADIEETPYHGSDFGKGBITMISYADEBEEAIVTWNIDLPY 180	Qy	494 -----
Qy	181 VDVVSNEREMIKRFLPQVLTREKPDVLTITYNGDNFDLPLYLKRAEKLGVTLLGRDKHPE 240	Db	1261 VFLLNSLGVSISIKGFDGSGVYRVYINEDLPFLQTFSRQKNTTYPNLIKEVLEEEIFGRKFQ 1320
Db	181 VDVVSNEREMIKRFLPQVLTREKPDVLTITYNGDNFDLPLYLKRAEKLGVTLLGRDKHPE 240	Qy	494 -----
Qy	241 PKHMRMGSFAVEIKGHLHDFLPVURTINLPTYLTEAIVLGKTKSKLGAEEIAAI 300	Db	1321 KNITFEKPKELADSKGKJDKRKYKLLDPLLNGDIVLDRVNTVKRETEGYYYDLSVEDNEN 1380
Db	241 PKHMRMGSFAVEIKGHLHDFLPVURTINLPTYLTEAIVLGKTKSKLGAEEIAAI 300	Qy	494 -----
Qy	301 WETBESEMKLQASMEADARATYELGKEFFPMEAELKLIGOSWMDVRSSTGNLYEWYLL 360	Db	1381 FLYFGFLIYAHNSYGYNGYKPAWYSKECASESVTAWGRHYTEMITKEIIEKFGKLYA 1440
Db	301 WETBESEMKLQASMEADARATYELGKEFFPMEAELKLIGOSWMDVRSSTGNLYEWYLL 360	Qy	543 D-----
Qy	361 RVAYERNELAQNPKDEEFYRRRLRTTLLGGYKTEPEGLWNTITYLDFRC----- 410	Db	1441 DSVTGDTEILVRNGRTRIEFPVIEKFLFVRYDIGKEYCILEDVAEALTLDRGKLIWKRKV 1500
Db	361 RVAYERNELAQNPKDEEFYRRRLRTTLLGGYKTEPEGLWNTITYLDFRC----- 420	Qy	544 -----
Qy	411 -----	Db	1501 PYMRHRRAKCKYRIVLNTNSWYDVTEDBSLIVADEGIGKEARPMETEGKSJLATAKDDSL 1560
Db	421 GKGIVNNSDVKEGDIYIIGDGNRQVKWVKHYEGKLJININGLKCJPNHKPVVTTENDRQ 480	Qy	544 -----
Qy	411 -----	Db	1561 VEYKPHAAEISYNGYYDIEVEGTHRFANGILVNTDGYAT1PGEKPKETKCKAKE 1620
Db	481 TRIRDSTAKSFSLCKVKSKIIITKLFFKXIAEFEKPKNPKSEEELTKGELSGIIAEGTLRK 540	Qy	566 FLKINTSKPLGLLIELEYFGFYLRGFVAKCRAVTDEGGRITRGLEVRRWSEJAKET 625
Qy	411 -----	Db	1621 FLKINTSKPLGLLIELEYFGFYLRGFVAKCRAVTDEGGRITRGLEVRRWSEJAKET 1680
Db	541 DIEYFDSSRGKKRISQYRVEITGENEKEELLRLYTFDKLFGIRPSVKRGKDNTALKI 600	Qy	626 QAKVLEAIIKEDSTKAEVIVDVEEAKYQVPLEXLVHQTQDSEYKAIQGHVHAI 685
Qy	411 -----	Db	1681 QAKVLEAIIKEDSTKAEVIVDVEEAKYQVPLEXLVHQTQDSEYKAIQGHVHAI 1740
Db	601 TTAKKAVYIQLIEELLNKIESLYAPAVLGEFFERDATYKIRSTIVVTOGTNNKWKIDIVA 660	Qy	686 AKRLAAKGKIKVRPSTIISIVLFRGSGKISDRVLLSEYDPKHKYDPYTYTENQVLPVAL 745
Qy	411 -----	Db	1741 AKRLAAKGKIKVRPSTIISIVLFRGSGKISDRVLLSEYDPKHKYDPYTYTENQVLPVAL 1800
Db	661 KLLDSLIGIPSYRSYKXIEENGKELTQHILEITGRDGLLFLQTLVGFISSEKNEALEKAE 720	Qy	746 RILAEFGYKEDLYQSSQKQVGLDNLK 774
Qy	411 -----	Db	1801 RILAEFGYKEDLYQSSQKQVGLDNLK 1829
Db	721 VREMRNLKNSFYNLSTFEVSSEYYKGEVYDITLEGGNPYFANGLIDHNSLYPSIIT 780	Qy	421 VSPDTLREGCKNYDVAPIVGYKFFCKDFFGFFPSPSTLGELITMRQEIKKMKATIDPIEKK 480
Qy	421 VSPDTLREGCKNYDVAPIVGYKFFCKDFFGFFPSPSTLGELITMRQEIKKMKATIDPIEKK 480	RESULT 2	DPOL_THELI

ID	DPOL	THELI	STANDARD;	PRT;	1702	AA.
AC	P30377;					
DT	01-APR-1993	(Rel. 25, Created)				
DT	01-APR-1993	(Rel. 41, Last sequence update)				
DT	28-FEB-2003	(Rel. 41, Last annotation update)				
DE	DNA Polymerase (EC 2.7.7.7) (Vent DNA Polymerase) [Contains: Endonuclease PI-TIIII (EC 3.1.1.-) (TII pol-1 intein) (IVPS1); DE Endonuclease PI-TIIII (EC 3.1.1.-) (TII pol-2 intein) (IVPS1)].					
GN	Thermococcus litoralis.					
OC	Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococaceae; Thermoococcus					
OC	NCBI_TAXID=2265;					
RN	[1]					
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.					
RX	MEDLINE=92302285; PubMed=1608969;					
RA	Perler F.B., Comb D.G., Jack W.E., Moran L.S., Qiang B.,					
RA	Carlow C.K.S., Jannasch H.,					
RA	"Intervening sequences in an Archaea DNA Polymerase Gene.";					
RL	Proc. Natl. Acad. Sci. U.S.A. 89:5577-5581(1992).					
RN	[2]					
RP	PROTEIN SPlicing.					
RX	Medline=3117083; PubMed=1415179;					
RA	Hodges R.A., Perler F.B., Noren C.J., Jack W.E.;					
RT	"Protein splicing removes intervening sequences in an archaea DNA					
RT	polymerase."					
RL	Nucleic Acids Res. 20:6153-6157(1992).					
CC	-!- FUNCTION: IN ADDITION TO POLYMERASE ACTIVITY, THIS DNA POLYMERASE EXHIBITS 3' TO 5' EXONUCLEASE ACTIVITY.					
CC	-!- FUNCTION: INTEIN ENCODED ENDONUCLEASES ARE THOUGHT TO MEDIATE SITE-SPECIFIC RECOMBINATION INITIATED BY ENDONUCLEASE CLEAVAGE AT THE "HOMING SITE" IN GENE THAT LACK THE INTEIN.					
CC	-!- CATALYTIC ACTIVITY: N deoxyribonucleoside triphosphate = N diphosphate + (DNA) (N).					
CC	-!- PTM: THIS PROTEIN UNDERGOES A PROTEIN SELF SPlicing THAT INVOLVES A POST-TRANSLATIONAL EXCISION OF THE TWO INTERVENING REGION (INTEINS), FOLLOWED BY PEPTIDE LIGATION.					
CC	-!- BIOTECHNOLOGY: Used in the PCR method because of its high thermostability and low error rate. Sold by New England Biolabs.					
CC	-!- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-B FAMILY.					
CC	-!- SIMILARITY: IN THE INTEIN SECTION? BELONGS TO THE HOMING ENDONUCLEASE FAMILY.					
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).					
CC	EMBL; M74198; AAA72101.1; -					
DR	PIR; S42459; S42459.					
DR	HSSP; P56689; 1TGO.					
DR	REBASE; 2613; PI-TIIII.					
DR	InterPro; IPR006172; DNA_Pol_B.					
DR	InterPro; IPR006134; DNA_Pol_B_dom.					
DR	InterPro; IPR006133; DNA_Pol_B_exo.					
DR	InterPro; IPR003586; Hedgehog_HintC.					
DR	InterPro; IPR003587; Hedgehog_HintN.					
DR	InterPro; IPR006141; Intein.					
DR	InterPro; IPR006142; INTEIN.					
DR	InterPro; IPR004576; Pol2.					
DR	Pfam; PF00136; DNA_Pol_B.					
DR	PRINTS; PR003104; DNA_Pol_B_exo.					
DR	SMART; SN00305; HintC.					
DR	SMART; SN00306; HintN.					
DR	SVRNGNKRFLYLMFNEIKDFISYPQKLEELWKGITLNGPRNCILKVDDEFQKLGY					
DR	SMART; SMO0486; POLBC.	1.				
DR	TIGRFAMS; TIGR01443; intein_Cterm.	2.				
DR	TIGRFAMS; TIGR01445; intein_Nterm.	2.				
DR	TIGRFAMS; TIGR00592; PC02.	2.				
DR	PROSITE; PS00116; DNA POLYMERASE_B.	FALSE_NEG.				
DR	PROSITE; PS50018; INTEIN_C.	2.				
DR	PROSITE; PS50019; INTEIN_N.	2.				
DR	PROSITE; PS50017; INTEIN_N_TER.	2.				
KW	Transferase; DNA-directed DNA polymerase; DNA replication;					
KW	DNA-binding; Hydrolase; Endonuclease; Autocatalytic cleavage;					
KW	Protein splicing; Intron homing.					
FT	CHAIN	1	494		DNA POLYMERASE.	1ST PART.
FT	CHAIN	1	1032		ENDONUCLEASE PI-TLI.	1.
FT	CHAIN	1	1081		DNA POLYMERASE.	2ND PART.
FT	CHAIN	1	1471		ENDONUCLEASE PI-TLI.	1.
FT	CHAIN	1	1702	1702	MW: 2165987C75F53B0	CR04.
SQ	SEQUENCE	1702	AA:	197293		
Query Match		83.4%		Score 3356;	DB 1;	Length 1702;
Best Local Similarity		42.7%		Pred. No. 8.7e-175;		
Matches		726;	Conservative	33;	Missmatches 15;	
Proc.		Indels 928;	Gaps	2;		
Qy	1	MIFDDYIITKGKPIRIFKENGKFIEUDPHFQPYYALLQDSDAIDBIKAIKGHRG	60			
Db	1	MILDTDYIITKGKPIRIFKENGKFIEUDPHFQPYYALLQDSDAIDBIKAIKGHRG	60			
Qy	61	KIVRTVDAYKVKKEPLGRDVEVWKLIIFEHPIFQDVALGKIREPAVVDIYEYDIPFARRY	120			
Db	61	KTVRLDAVYKVKKEPLGRDVEVWKLIIFEPDQVAMRGKIREPAVVDIYEYDIPFARRY	120			
Qy	121	LIDKGKLMPEGEDEEKLMAFDIETYHEGDEFKGKEIIMSYADEEARVTKWKNIDLY	180			
Db	121	LIDKGKLMPEGEDEEKLMAFDIETYHEGDEFKGKEIIMSYADEEARVTKWKNIDLY	180			
Qy	181	YDVVSNERENIKRFQIVREKDPDVLITYGDNEDLPYLIKRAEKLGYTLLGRDEKEHP	240			
Db	181	VDVSNERENIKRFQIVREKDPDVLITYGDNEDLPYLIKRAEKLGYTLLGRDEKEHP	240			
Qy	241	PKIHRNGDSFAVEIKGRHEDLFPVTRRNLPPYTLLEAVYEAVLGKTKSKLGAEEIAAI	300			
Db	241	PKIQRNGDSFAVEIKGRHEDLFPVTRRNLPPYTLLEAVYEAVLGKTKSKLGAEEIAAI	300			
Qy	301	WETBEMKCLLAQYSMEDARATYELGKEPPMBAEAKLIGQSYWDSSSTGMLVEWILL	360			
Db	301	WETTEEMKCLLAQYSMEDARATYELGKEPPMBAEAKLIGQSYWDSSSTGMLVEWILL	360			
Qy	361	RVAYERNELLAQPKPDEEYRRRLTYLGYYKEPGLYNTDGRCLYPSITIVTHN	420			
Db	361	RVAYRNELLAPKPKPDEEYRRRLTYLGYYKEPGLYNTDGRCLYPSITIVTHN	420			
Qy	421	VSPDTLREGCKNYDAPIVGYKFCKDFFGFPISPSILGELITWQEKCKMKATIDPIEKK	480			
Db	421	VSPDLEKEGKNTIVAPIVGYRFKDFGGFPISPSILGDLIAMEQDIIKGMKSTIDPIEKK	480			
Qy	481	MLDYRORAYKLHA-----	493			
Db	481	MLDYRORAYKLHA-----	493			
Qy	541	VNNLIAFSNKIKKESEVKVKALRHKTGKAYEIQSSGRKINITAGHSLFVNRGHI	600			
Qy	494	-----	493			
Db	601	KEVSGDGIKEGDLIVTAPKKKLNEKGVSINTIPELISLSEEETADIVMTISAKGRKNPFK	660			
Qy	494	-----	493			
DR	SVRNGNKRFLYLMFNEIKDFISYPQKLEELWKGITLNGPRNCILKVDDEFQKLGY	780				

Qy	494	-----	493	-----	DB DNA polymerase (EC 2.7.7.7) (TO POL).
Db	781	VSEGYAAQQRNKTGGISYSVLYNEDNVLESMRNVALEKFFGKVRVDRNCVSISSKRMAYL	840	-----	GN OS Thermococcus gorgonarius.
Qy	494	-----	493	-----	OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae;
Qy	494	-----	493	-----	OC Thermococcus
Db	B 41	VMKCLCGALAENKRIPPSVLTSPEPVWSFLEAYFTGDTHPSKRFLSTKSELLANQL	900	-----	NCBI_TaxID=71997;
Qy	494	-----	493	-----	RN [1] X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
Db	901	VFLINSLGISSYKIGFDGKVRYVINYDLOFFQTSREKNTTYSNLPIKEIIRDVFGKEFQ	960	-----	RX MEDLINE-99192230; PubMed=10097083;
Qy	494	-----	493	-----	RA Hopfner K.-P., Eichinger A., Engle R.A., Luew F., Ankenbauer W.,
Db	961	KMTTFKKFKEVYDVGKLNREKAKLLEFFINGDIVLDRKSYREKDYGGYDLSVDEDNEN	1020	-----	RA Huber R., Angerer B.,
Qy	494	-----	543	-----	RT "Crystal structure of a thermostable type B DNA polymerase from Thermococcus gorgonarius".
Db	1021	FLVFGFLIYAHNSYGGMGPARYSKCAESVTAGMGRHYEMTIREBIEEKFGPKVLYA	1080	-----	RT Proc. Natl. Acad. Sci. U.S.A. 96:3600-3605 (1999).
Qy	543	D-----	543	-----	RL -1- FUNCTION: IN ADDITION TO POLYMERASE ACTIVITY, THIS DNA POLYMERASE EXHIBITS 3' TO 5' EXONUCLEASE ACTIVITY.
Db	1081	DVSYGSGESEIIIRQNGKRFVKTIDFLPKVDSIGERKEYCILEGVVEATLDDDGKLYWKPV	1140	-----	CC CC -1- CATALYTIC ACTIVITY: N deoxyribonucleoside triphosphate = N diphosphate + [DNA] (N).
Qy	544	-----	543	-----	CC CC -1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-B FAMILY.
Db	1141	PTVNRHANKRMFRILWNTNSWYIDVTDHSLIGYLNSTSKTAKKIGERLKEVKFPELKG	1200	-----	CC CC DR PDB: 1TGO; 22-MAR-99.
Qy	544	-----	543	-----	DR InterPro: IPRO06172; DNA_Pol_B.
Db	1201	AVKSLICPNAPLDENTTKTSETIAVKFWELVGLIVGDDNWGGDSRWAELYGLSTGXDAE	1260	-----	DR InterPro: IPRO06134; DNA_Pol_B_dom.
Qy	544	-----	543	-----	DR InterPro: IPRO06133; DNA_Pol_B_exo.
Db	1261	IKQXLLIEPLKTYGVISNTYPKNEKGDDNIALAKSLVKMFKDFDKRKRKPEFMYLPV	1320	-----	DR InterPro: IPRO04578; PolB.
Qy	544	-----	543	-----	DR Pfam: PF00136; DNA_Pol_B_1.
Db	1321	TYIEAFLRGLFSAADGTVTIRKGVPFIRLTNTDADFLREVRLWIVGINSIFAETTPNR	1380	-----	DR Pfam: PF03104; DNA_Pol_B_exo; 1.
Qy	544	-----	543	-----	DR PRANTS; PRO0106; DNAPOLB.
Db	1381	YNGVSTGTYSKHLRINKWKRFAERIGFLIERKQKRLLLEHLSARVKRNTIDFGFDLVHVK	1440	-----	DR SMART; SM00466; POLB; 1.
Qy	544	-----	543	-----	DR PROSITE; PS00116; DNA_Polymerase_B; 1.
Db	1441	KVEEPIPYEGVYDIEVEETHRFANNLVNNDGFATIPGKPKPLKKRKAKEFLYINS	1500	-----	DR TRANSFERASE; DNA-directed DNA Polymerase; DNA replication;
Qy	573	KLPQSLLELEYEGFLRGGFVKARYAVIDEDEGRITGGLEVRRDSDSEIAKETOAKVLEA	632	-----	KW KW DNA-binding; Hydrolase; Nuclease; Exonuclease; Multifunctional enzyme;
Db	1501	KLPQSLLELEYEGFLRGGFVKARYAVIDEDEGRITOGLEVRRDSDSEIAKETOAKVLEA	1560	-----	KW KW 3D-STRUCTURE
Qy	633	TIKEDSTFKEAVEIYKDVVEEIAKCYVPLEKVIWHEQTKDLSYEKAIGPHAJAKLAAK	692	-----	KW KW DISULFID 428 442
Db	1561	ILKGSVEGAEVDRDVVEKIAKYRVPFLBKLVHEQTRDLDYKAGPHAJAKLAA	1620	-----	FT FT DISULFID 506 509
Qy	693	GIKVRPGTIIISYIYLRGSKISDRVLLSEYDPKGKHYDPDYYIENGVLPAVTRILEAFG	752	-----	FT FT STRAND 2 10
Db	1621	GIKVRPGTIIISYIYLGSKISDRVLLSEYDPKGKHYDPDYYIENGVLPAVTRILEAFG	1680	-----	FT FT STRAND 11 12
Qy	753	YRKEDLKYQSSKQVGLDAWIKK	774	-----	FT FT STRAND 13 22
Qy	1681	YRKEDLRYQSSKQVGLDAWIKR	1702	-----	FT FT STRAND 23 24
Db	-----	-----	-----	-----	FT FT STRAND 25 31
Qy	-----	-----	-----	-----	FT FT STRAND 37 42
Db	-----	-----	-----	-----	FT FT DISULFID 506 509
Qy	-----	-----	-----	-----	FT FT STRAND 52 53
Db	-----	-----	-----	-----	FT FT STRAND 53 58
Qy	-----	-----	-----	-----	FT FT STRAND 55 60
Db	-----	-----	-----	-----	FT FT STRAND 61 64
Qy	-----	-----	-----	-----	FT FT STRAND 67 75
Db	-----	-----	-----	-----	FT FT HELIX 45 47
Qy	-----	-----	-----	-----	FT FT HELIX 48 51
Db	-----	-----	-----	-----	FT FT STRAND 52 53
Qy	-----	-----	-----	-----	FT FT STRAND 53 58
Db	-----	-----	-----	-----	FT FT STRAND 55 60
Qy	-----	-----	-----	-----	FT FT STRAND 61 64
Db	-----	-----	-----	-----	FT FT STRAND 67 75
Qy	-----	-----	-----	-----	FT FT HELIX 76 77
Db	-----	-----	-----	-----	FT FT TURN 76 77
Qy	-----	-----	-----	-----	FT FT TURN 77 78
Db	-----	-----	-----	-----	FT FT TURN 78 86
Qy	-----	-----	-----	-----	FT FT TURN 90 91
Db	-----	-----	-----	-----	FT FT TURN 91 102
Qy	-----	-----	-----	-----	FT FT TURN 102 104
Db	-----	-----	-----	-----	FT FT TURN 104 105
Qy	-----	-----	-----	-----	FT FT TURN 105 106
Db	-----	-----	-----	-----	FT FT TURN 106 110
Qy	-----	-----	-----	-----	FT FT TURN 110 116
Db	-----	-----	-----	-----	FT FT TURN 116 123
Qy	-----	-----	-----	-----	FT FT TURN 124 125
Db	-----	-----	-----	-----	FT FT STRAND 125 137
Qy	-----	-----	-----	-----	FT FT STRAND 137 144
Db	-----	-----	-----	-----	FT FT TURN 144 152
Qy	-----	-----	-----	-----	FT FT STRAND 152 153
Db	-----	-----	-----	-----	FT FT TURN 153 157
Qy	-----	-----	-----	-----	FT FT TURN 157 164
Db	-----	-----	-----	-----	FT FT TURN 164 166
Qy	-----	-----	-----	-----	FT FT TURN 166 166
Db	-----	-----	-----	-----	FT FT TURN 166 166
Qy	-----	-----	-----	-----	FT FT HELIX 211 213
Db	-----	-----	-----	-----	FT FT HELIX 214 214
Qy	-----	-----	-----	-----	FT FT HELIX 215 225
Db	-----	-----	-----	-----	FT FT TURN 226 227
Qy	-----	-----	-----	-----	FT FT TURN 227 235
Db	-----	-----	-----	-----	FT FT STRAND 234 243
Qy	-----	-----	-----	-----	FT FT STRAND 240 240
Db	-----	-----	-----	-----	FT FT STRAND 243 251
Qy	-----	-----	-----	-----	FT FT STRAND 251 253
Db	-----	-----	-----	-----	FT FT TURN 253 253

RESULT 3
 DPOL_TheGO STANDARD; PRT; 773 AA.
 ID DPOL_TheGO
 AC P56619;
 DT 15-JUL-1999 (Rel: 38, Created)
 DT 15-JUL-1999 (Rel: 38, Last sequence update)
 DT 28-FEB-2003 (Rel: 41, Last annotation update)

Matches	Conservative	83; Mismatches	82; Indels	3; Gaps	2;
256	STRAND	259			
FT	HELIX	260			
FT	HELIX	267			
FT	HELIX	275	283		
FT	HELIX	292	301		
FT	TURN	303	304		
FT	HELIX	305	337		
FT	TURN	338	338		
FT	HELIX	341	345		
FT	TURN	346	346		
FT	HELIX	349	363		
FT	TURN	364	365		
FT	STRAND	366	366		
FT	HELIX	374	379		
FT	STRAND	389	390		
FT	STRAND	396	405		
FT	HELIX	408	415		
FT	TURN	416	417		
FT	STRAND	420	422		
FT	STRAND	423	423		
FT	TURN	424	424		
FT	TURN	426	427		
FT	STRAND	431	433		
FT	TURN	435	437		
FT	STRAND	440	442		
FT	HELIX	448	469		
FT	HELIX	473	490		
FT	TURN	491	492		
FT	HELIX	493	498		
FT	TURN	500	501		
FT	TURN	503	504		
FT	HELIX	507	530		
FT	TURN	531	532		
FT	STRAND	535	540		
FT	STRAND	543	547		
FT	TURN	549	550		
FT	HELIX	553	568		
FT	TURN	559	570		
FT	TURN	573	574		
FT	STRAND	577	590		
FT	TURN	591	592		
FT	STRAND	593	597		
FT	TURN	599	600		
FT	STRAND	603	606		
FT	STRAND	615	615		
FT	HELIX	617	631		
FT	TURN	632	633		
FT	HELIX	635	651		
FT	TURN	652	653		
FT	HELIX	657	659		
FT	STRAND	662	665		
FT	HELIX	670	672		
FT	HELIX	679	689		
FT	TURN	690	690		
FT	STRAND	693	704		
FT	TURN	710	711		
FT	STRAND	714	716		
FT	HELIX	717	719		
FT	TURN	722	724		
FT	STRAND	727	727		
FT	HELIX	729	735		
FT	TURN	736	737		
FT	HELIX	738	740		
FT	TURN	741	741		
FT	HELIX	742	746		
FT	TURN	747	749		
FT	HELIX	752	754		
FT	TURN	766	767		
FT	TURN	771	773		
SEQUENCE	773 AA:	89812 MW:	F67AF04E875FBE44	CRC64;	

"Cloning of thermostable DNA Polymerases from hyperthermophilic marine Archaea with emphasis on Thermococcus sp. 9ON-7 and mutations affecting 3', 5' exonuclease activity."

RT Proc. Natl. Acad. Sci. U.S.A. 93:5281-5285 (1996).

CC -!- CATALYTIC ACTIVITY: N deoxyribonucleoside triphosphate = N diphosphate

CC + [DNA] (N).

CC -!- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-B FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to licensee@isb-sib.ch).

CC

DR Pfam: PF03104; DNA pol B-exc; 1.

DR PRNTS; PRO016; DNAPOB.

DR SMART; SM00488; POLB_C; 1.

DR TIGRFAMS; TIGR00522; POL2; 1.

DR PROSITE; PS00116; DNA_POLYMERASE_B; 1.

DR TRANSFERASE; DNA-directed DNA polymerase; 1.

DR KWDNA-binding; Cmpd proteome.

FT VARIANT 532 R -> S (IN STRAIN GE21).

FT VARIANT 553 HE -> P (IN STRAIN GE23).

FT VARIANT 756 K -> R (IN STRAINS GE23 AND ST 855).

FT CONFLICT 263 V -> A (IN REF. 1).

FT CONFLICT 277 A -> T (IN REF. 1).

FT CONFLICT 281 A -> V (IN REF. 1).

FT CONFLICT 320 F -> S (IN REF. 1).

FT CONFLICT 339 Q -> H (IN REF. 1).

FT CONFLICT 359 R -> T (IN REF. 1).

FT CONFLICT 391 N -> N (IN REF. 1).

FT CONFLICT 771 AA; 83946 NW; 110A87045ABA5522 CRC64;

SEQ

Query Match 78.6%; Score 3164.5%; Length 771;

Best Local Similarity 76.1%; Prod. No. 8e-165; 2;

Matches 388; Conservative 89; Mismatches 93; Indels 3; Gaps 2;

Qy 1 MIFDPTDYYTIDGKPIIIRIFKKGENSEFKEFIELDPDPHQPYTALLKKDSAIIDKEIAGERHG 60

Db 1 MTDADYYTEDGKPIIIRIFKKGEEFKFVYDTRFPIYALLKKDSAIIDKEIAGERHG 60

Qy 61 KIVTVDAYKVKRKEFKLGRDVEWKLIFHPDQVPAFLRGKIKHEPAVIDIYHYDIPPAKRY 120

Db 61 KIVTIEFVQKCFPLGPVIEWKLYLHPQDVPATREKIRHPAVIDIYHYDIPPAKRY 120

Qy 121 LIDKGLIIMEGDDEKLMAFDIETFYHEGDERGKGELEMISYADEEARVITWQDIDPY 180

Db 121 LIDKGLIIMEGDDEKLMAFDIETFYHEGDERGKGELEMISYADEEARVITWQDIDPY 180

Qy 181 VDVSNSNREMIKEFQV1TREKDVLITYNGNFDLDPYLIRKAEKGUTVLLGRDKHEPE 240

Db 181 VEVSSSEFEMIKLVKTVREKDVLITYNGNFDFFYLIRKAEKGULPLGRD-NSE 238

Qy 241 PKHRMGDSFAV1KGRHFDFPVERTINPPTYLEAVTAVALTKSKLIGEETIAI 300

Db 239 PKNCRMGSVLAEB1KGRHFDFPVIRTTINPPTYLEAVTAIFGSKKERVYAB1EA 298

Qy 301 WEPBE5MKLQAQSMDERATYELKEF1FPMMAELAKLIGGSVWDVSRSSTGONLYEWYLL 360

Db 299 WEGKGLERPVARYSMSMDERATYELKEF1FPMMAELAKLIGGSVWDVSRSSTGONLYEWYLL 358

Qy 361 RAYERNEELAPNPKPDEEFYRRLRTTLYGGYKKEPERGLWNTYLDFLCPLYSP1VTHN 420

Db 359 RKAVERNEELAPNPKPDERYERIRESFEGGYREPEGLWNGIVSDFLRSYPS1I1H 418

Qy 421 VSPDTLBRREGKQYDVAFLVGGYKFCKFQGPGETPSLIGELIMPOETKKKKATDIEKK 480

Db 419 VSPDTLBRREGKQYDVAFLVGGYKFCKFQGPGETPSLIGELIMPOETKKKKATDIEKK 478

Qy 481 MLDYRQRAVGLHANSYCGYMGYPKARYSKCEAESTAAGRYH1EMTKE1EKKGFKVL 540

Db 479 LLDDYRQRAVGLHANSYCGYMGYPKARYSKCEAESTAAGRYH1EMTKE1EKKGFKVL 537

Qy 541 YADTDGFYATIPEKPKTIIKAKFELKFKY1NSKLPG1LLEYEFGYLRLGK1YV1RSGSK1SDR1V1L 597

Db 538 Y1PTDGLYATIPEKPKTIIKAKFELKFKY1NSKLPG1LLEYEFGYLRLGK1YV1RSGSK1SDR1V1L 597

Qy 601 DEGRTTIGLERYVRDWE1A1KETQAVKVLB1LKE1DSDVCAE1VYKQVPL 660

Db 598 DEEGK1VTRGL1EVRDWE1A1KETQAVKVLB1LKE1DSDVCAE1VYKQVPL 657

Qy 661 EKLV1THE101KDLSEYK1A1KIGPHVA1A1KRLAAGK1YV1RSGSK1SDR1V1L 720

Db 658 EKLV1YEQ1TRP1SEYK1A1KIGPHVA1A1KRLAAGK1YV1RSGSK1SDR1V1L 717

Qy 721 SEYDPKK1KVDYDXTEN0V1PAVL11EAFGKRD1LYQKQSSKQVGLDAW1K 773

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

DR 718 EEFDDPKK1KVDYDXTEN0V1PAVL11EAFGKRD1LYQKQSSKQVGLDAW1K 770

DR PROSITE; PRO016; DNAPOB.

DR SMART; SM00488; POLB_C; 1.

DR TIGRFAMS; TIGR00522; POL2; 1.

DR PROSITE; PS00116; DNA_POLYMERASE_B; 1.

DR TRANSFERASE; DNA-directed DNA polymerase; 1.

DR KWDNA-binding; Cmpd proteome.

FT VARIANT 532 R -> S (IN STRAIN GE21).

FT VARIANT 553 HE -> P (IN STRAIN GE23).

FT VARIANT 756 K -> R (IN STRAINS GE23 AND ST 855).

FT CONFLICT 263 V -> A (IN REF. 1).

FT CONFLICT 277 A -> T (IN REF. 1).

FT CONFLICT 281 A -> V (IN REF. 1).

FT CONFLICT 320 F -> S (IN REF. 1).

FT CONFLICT 339 Q -> H (IN REF. 1).

FT CONFLICT 359 R -> T (IN REF. 1).

FT CONFLICT 391 N -> N (IN REF. 1).

FT CONFLICT 771 AA; 83946 NW; 110A87045ABA5522 CRC64;

SEQ

Query Match 78.6%; Score 3164.5%; Length 771;

Best Local Similarity 76.1%; Prod. No. 8e-165; 2;

Matches 388; Conservative 89; Mismatches 93; Indels 3; Gaps 2;

Qy 1 MIFDPTDYYTIDGKPIIIRIFKKGENSEFKEFIELDPDPHQPYTALLKKDSAIIDKEIAGERHG 60

Db 1 MTDADYYTEDGKPIIIRIFKKGEEFKFVYDTRFPIYALLKKDSAIIDKEIAGERHG 60

Qy 61 KIVTVDAYKVKRKEFKLGRDVEWKLIFHPDQVPAFLRGKIKHEPAVIDIYHYDIPPAKRY 120

Db 61 KIVTIEFVQKCFPLGPVIEWKLYLHPQDVPATREKIRHPAVIDIYHYDIPPAKRY 120

Qy 121 LIDKGLIIMEGDDEKLMAFDIETFYHEGDERGKGELEMISYADEEARVITWQDIDPY 180

Db 121 LIDKGLIIMEGDDEKLMAFDIETFYHEGDERGKGELEMISYADEEARVITWQDIDPY 180

Qy 181 VDVSNSNREMIKEFQV1TREKDVLITYNGNFDLDPYLIRKAEKGUTVLLGRDKHEPE 240

Db 181 VEVSSSEFEMIKLVKTVREKDVLITYNGNFDFFYLIRKAEKGULPLGRD-NSE 238

Qy 241 PKHRMGDSFAV1KGRHFDFPVERTINPPTYLEAVTAVALTKSKLIGEETIAI 300

Db 239 PKNCRMGSVLAEB1KGRHFDFPVIRTTINPPTYLEAVTAIFGSKKERVYAB1EA 298

Qy 301 WEPBE5MKLQAQSMDERATYELKEF1FPMMAELAKLIGGSVWDVSRSSTGONLYEWYLL 360

Db 299 WEGKGLERPVARYSMSMDERATYELKEF1FPMMAELAKLIGGSVWDVSRSSTGONLYEWYLL 358

Qy 361 RAYERNEELAPNPKPDEEFYRRLRTTLYGGYKKEPERGLWNTYLDFLCPLYSP1VTHN 420

Db 359 RKAVERNEELAPNPKPDERYERIRESFEGGYREPEGLWNGIVSDFLRSYPS1I1H 418

Qy 421 VSPDTLBRREGKQYDVAFLVGGYKFCKFQGPGETPSLIGELIMPOETKKKKATDIEKK 480

Db 419 VSPDTLBRREGKQYDVAFLVGGYKFCKFQGPGETPSLIGELIMPOETKKKKATDIEKK 478

Qy 481 MLDYRQRAVGLHANSYCGYMGYPKARYSKCEAESTAAGRYH1EMTKE1EKKGFKVL 540

Db 479 LLDDYRQRAVGLHANSYCGYMGYPKARYSKCEAESTAAGRYH1EMTKE1EKKGFKVL 537

Qy 541 YADTDGFYATIPEKPKTIIKAKFELKFKY1NSKLPG1LLEYEFGYLRLGK1YV1RSGSK1SDR1V1L 597

Db 538 Y1PTDGLYATIPEKPKTIIKAKFELKFKY1NSKLPG1LLEYEFGYLRLGK1YV1RSGSK1SDR1V1L 597

Qy 601 DEGRTTIGLERYVRDWE1A1KETQAVKVLB1LKE1DSDVCAE1VYKQVPL 660

Db 598 DEEGK1VTRGL1EVRDWE1A1KETQAVKVLB1LKE1DSDVCAE1VYKQVPL 657

Qy 661 EKLV1THE101KDLSEYK1A1KIGPHVA1A1KRLAAGK1YV1RSGSK1SDR1V1L 720

Db 658 EKLV1YEQ1TRP1SEYK1A1KIGPHVA1A1KRLAAGK1YV1RSGSK1SDR1V1L 717

Qy 721 SEYDPKK1KVDYDXTEN0V1PAVL11EAFGKRD1LYQKQSSKQVGLDAW1K 773

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

CC :||||| :||||| :||||| :||||| :|||||

between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

RESULT 7
DEPOL_PYRSD STANDARD; PRT; 1312 AA.
ID DEPOL_PYRSD; PRT; 1312 AA.
AC Q01334; Q51335; Q51336; Created
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE DNA polymerase [EC 2.7.7.1. (Deep vent DNA polymerase) [Contains:
DE Endonucleases PI_PSP1 [EC 3.1.1.1. (Psp-GDB pol intein)].
GN POL.
OS Pyrococcus sp. (strain GB-D).
OC Archaea: Buryarchaeota; Thermococci; Thermococcaceae;
OC Pyrococcus
RN [NCBI_TaxID:69013;]

RP SEQUENCE FROM N.A., AND SEQUENCE OF 493-517.
 RX MEDLINE=94094330; PubMed=865515;
 RA XU M.-Q.; Southworth M.W.; Marsha F.B.; Horstra L.J.; Parler P.B.;
 RT "In vitro protein splicing of purified precursor and the
 RT identification of a branched intermediate.",
 RL Cell 75:1371-1377 (1993).
 CC -1- FUNCTION: IN ADDITION TO POLYMERASE ACTIVITY, THIS DNA POLYMERASE
 CC EXHIBITS 3' TO 5' EXONUCLEASE ACTIVITY.
 CC -1- FUNCTION: INTEIN ENCODED ENDONUCLEASES ARE THOUGHT TO MEDIATE
 CC A POST-TRANSLATIONAL EXCISION OF THE INTERVENING REGION (INTEIN)
 CC FOLLOWED BY PEPTIDE LIGATION.
 CC -1- BIOTECHNOLOGY: Used in the PCR method because of its high
 CC THERMOSTABILITY and low error rate. Sold by New England Biolabs.
 CC -1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-B FAMILY.
 CC -1- SIMILARITY: IN THE INTEIN SECTION, BELONGS TO THE HOMING
 CC ENDONUCLEASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires license agreement (See <http://www.isb-sib.ch/announce>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; U00707; AAA67130.1; -.
 DR EMBL; U00707; AAA67131.1; -.
 DR EMBL; U00707; AAA67132.1; -.
 DR PIR; S68593; S68593.
 DR HSSP; P5680; 1100.
 DR REBASE; 261; PI-BSPI.
 DR InterPro; IPR006172; DNA_Pol_B.
 DR InterPro; IPR006134; DNA_Pol_B_dom.
 DR InterPro; IPR006133; DNA_Pol_B_exo.
 DR InterPro; IPR003586; Hedgehog_hntN.
 DR InterPro; IPR003587; Hedgehog_hntN.
 DR InterPro; IPR006141; Intein.
 DR InterPro; IPR006142; Intein.
 DR InterPro; IPR004042; Intein_endonuc.
 DR InterPro; IPR004578; Pol2.
 DR Pfam; PF003104; DNA_Pol_B_exo; 1.
 DR Pfam; PF003104; DNA_Pol_B_exo; 2.

DR	PRINTS	PR00106; DNAPOLB.	Db	719	VRYNQNKREYLVEFNAYRDVLSIMPEEBELKEWKRIGTRNGTFRMGTVDIDEDFAKLGYVV	778
DR	SMART	PR00379; INTEIN.	Qy	494		
DR	SMART	SM00305; HintC; 1.	Db	779	SEGSARKKWNQTCGWSYTFLYNENDVLDMEHLAKKPFKVKRGKNNYVEPKMKMAYII	838
DR	SMART	SM00306; HintN; 1.	Qy	494		
DR	PROSITE	PS00486; POLBC; 1.	Db	839	FESLQGTLAENKRVPEVFTSSKGVRWAFLEGYFIGDGDVHPSKRVLSTKSLLVNLG	898
DR	PROSITE	TIGR01443; interin_Cterm; 1.	Qy	494		
DR	PROSITE	TIGR01445; interin_Cterm; 1.	Db	899	LLNLSLGWSA1KGYDGSYVYRVVNEELKFTEYRKKNVYHSH1VPIKDILKETFGKVQK	958
DR	PROSITE	TIGR01447; interin_Nterm; 1.	Qy	494		
DR	PROSITE	PS00116; DNA POLYMERASE_B; 1.	Db	959	NISYKFRFELVNGRLDREKARRIIEWLNGDIVLDRVVEIRREYDGYVYDLSVDEDNF	1018
DR	PROSITE	PS00818; INTEIN_CTER; 1.	Qy	494		
DR	PROSITE	PS00819; INTEIN-ENDONUCLEASE; 1.	Db	1019	LAGFGFLYAHNSYYGGYAKARWYCEAESTWAGREYIFVRLKEEKGFKVLYAD	1078
DR	TRANSF	TIGR01487; interin_N_TER; 1.	Qy	494		
KW	Transf erase	DNA-directed DNA polymerase; DNA replication;	Db	1079	TDGLYATIPGAKPEEKVKALEYFDVYDINAFLKPLLELEYEGFVTRGFFVTKCYALIDEE	1138
KW	DNA-binding	Hydrolase; Endonuclease; Autocatalytic cleavage;	Qy	544	TDGTYATIPGKEKPTIKKAKAELFLKLYINSKLPOLLELEYEGFYLGRPFVAKRYAVDEE	603
KW	Protein splicing	Protein splicing; Inttron homing.	Db	1079	TDGLYATIPGAKPEEKVKALEYFDVYDINAFLKPLLELEYEGFVTRGFFVTKCYALIDEE	1138
FT	CHAIN	1 493 1029 DNA POLYMERASE_1ST PART.	Qy	604	GRITTRGLEVVRDWWSEAKETAKVLEAILEDSEVKADEVKAVLWVVEIAKVQVPLDEKL	663
FT	CHAIN	1030 1312 DNA POLYMERASE_2ND PART.	Db	1139	GKIIITRGLEVVRDWWSEAKETAKVLEAILEDSEVKADEVKAVLWVVEIAKVQVPLDEKL	1198
SQ	SEQUENCE	1312 AA; 152852 MW; B6751880564ID26A CRC64;	Qy	664	VIHQITKDSEYKAIGPHVATKRLAAGKIKVPRGTTISYVLRGSKSISDRIVILSEY	723
Qy	Query Match	72.9% Score 2934.5; DB 1; Length 1312;	Db	1199	VIVSQITRPLHEYKAIGPHVAVKRLAARGKVTPGMWIGYVLRGDPISKRALLAEF	1258
Matches	Best Local Similarity	45.6% Pred. No. 5e-152;	Qy	724	DPKKHPDYYIENQVLPAVRLAEFGYRKEKDLYKQSSKQVGLDAWL	772
Matches	597; Conservative	86; Mismatches 87; Indels 539; Gaps 3;	Db	1259	DLRKHYDAEYIENQVLPAVRLAEFGYRKEKDLYKQSSKQVGLDAWL	1307
Qy	1	MIPDTYITKDGKPIIRFKKENGEFKBILDHFQPYIYALLKKDDAIDETKAIKGHRG 60	RESULT 8			
Db	1	MILDADYITEDGAPIIRFKKENGEFKYEDNFRPYIYALLKKDDSQIDEVKITAERHG 60	DPOL_PYRHO	STANDARD;	PRT;	1235 AA.
Qy	61	KIVRVDVAKVCKFLGRDIDVWLKLIPEHPQVPAIRGKREHPATVDTYDIPFAKRY 120	ID	DPOL_PYRHO		
Db	61	KIVRVDVAKVCKFLGRDIDVWLKLIPEHPQVPAIRGKREHPATVDTYDIPFAKRY 120	AC	OS97TO		
Qy	121	VDYVSNBREMIKRFVQIWRKEDVPLKFLYKFLVWRLYFHPDQVPAIRDKIREHSAVIDFYDIPFAKRY 120	DT	15-DBC-1998 (Rel. 37, Created)		
Db	121	VDYVSNBREMIKRFVQIWRKEDVPLKFLYKFLVWRLYFHPDQVPAIRDKIREHSAVIDFYDIPFAKRY 120	DT	15-DBC-1998 (Rel. 37, Last annotation update)		
Qy	181	PKTHRMGDSFAVBIKGKTHEDLFPVYRRTINLPPTYLEAVYEAULGKTKSKLGAEIAI 300	DE	28-FEB-2003 (Rel. 41, Last annotation update)		
Db	244	PKTHRMGDSFAVBIKGKTHEDLFPVYRRTINLPPTYLEAVYEAULGKTKSKLGAEIAI 300	DE	DNA Polymerase (EC 2.7.7.7) [Contains: Pho pol integrin (Pho Pol I DE integrin)].		
Qy	239	PKNQRLGIMTAVBIKGKTHEDLFVHTRRTINLPPTYLEAVYEAULGKTKSKLGAEIAI 298	DE	DE integrin (Pho Pol I DE integrin).		
Db	301	WEEFEESMKKLAQYSMEDRATYELGKEFPMBELAKLIGQSVWDVRSSTGNLVEWYLL 360	GN	POI_OR PH1947 OR PHBTO47.		
Qy	301	WEEFEESMKKLAQYSMEDRATYELGKEFPMBELAKLIGQSVWDVRSSTGNLVEWYLL 360	OS	Pyrococcus horikoshii.		
Db	299	WTGKGGLERVAKYSMEAKVYFBLGRFPEMAOLSRLVQPLWDVRSSTGNLVEWYLL 358	OC	Archaeota; Thermococcales; Thermococcales; Thermococcales;		
Qy	361	RVAYERNLAPNPKDEEEYRRRLRTTYGGTYKEPEKGKLGVLSDERSYPSIITIN 420	OC	Pyrococcus.		
Db	359	RKAYERNLAPNPKDEEEYRRRLRTTYGGTYKEPEKGKLGVLSDERSYPSIITIN 418	OX	Pyrococcus horikoshii.		
Qy	421	VSPDTLREBEGKTYDVAFPIGVKFCDFPGFIPSILCEGELITMROEIKKMKATDPIEKK 480	RN	SEQUENCE FROM N.A.		
Db	419	VSPDTLNEEGCREYDVAEFGHKFCDFPGFIPSILKRLLDERQEIKRKKASKDIEKK 478	RN	SEQUENCE FROM N.A.		
Qy	481	MUDYRQPKV-----490	RC	SEQUENCE FROM N.A.		
Db	479	MUDYRQPKV-----538	RX	SEQUENCE FROM N.A.		
Qy	491	---LHA-----493	RA	PubMed=96344137; Kawarabayasi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y., Nagai Y., Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoya A., Nakazawa H., Takamiya M., Ohfuku Y., Sakai M., Ogura K., Otsuka R., Nakazawa H., Kuhida N., Oguchi A., Funahashi T., Tanaka T., Kudo H., Yamazaki J., Kuhida N., Oguchi A., Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K., Masuchi Y., Shizuya H., Kuroki H.,		
Db	539	VAGTHAFSDRKSKKARYMAVAKVIRYSGNVRIVLNSGRKITTEGHSLFVTRNGDL 598	RA	"Complete sequence and gene organization of the genome of a hyper-		
Qy	494	-----493	RA	thermophilic archaebacterium, Pyrococcus horikoshii OR3.";		
Db	599	VEATGEDVYKIGDILLLAVPVSNLPEKERLNIVELLNLSPPEETDILTIPVKGRKNFK 658	RT	-CATALYTIC ACTIVITY: N deoxy nucleoside triphosphate = N diphosphate		
Qy	494	-----493	CC	+ (DNA) (N).		
Db	659	GMRLTRLWIFGEKVKRVTASRYLRLHNLGYIRLKIGGYDILDEGLEKTYEKLVDV 718	CC	- PTM: THIS PROTEIN UNDERGOES A PROTEIN SELF SPLICING THAT INVOLVES A POST-TRANSLATIONAL EXCISION OF THE INTERVENTING REGION (INTEIN)		
Qy	494	-----493	CC	- FOLLOWED BY PEPTIDE LIGATION (POLYMER).		
		-----	CC	- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-B FAMILY.		

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to licensee@isb-sib.ch).

EMBL; AP000007; BAA31074.1; -.

DR; PIR; C71210; C71210.

DR; HSSP; P56689; 1TG0.

DR; InterPro; IPRO01172; DNA_Pol_B.

DR; InterPro; IPRO06134; DNA_Pol_B_exo.

DR; InterPro; IPRO06133; DNA_Pol_B_dcm.

DR; InterPro; IPRO03586; Hedgehog_FintN.

DR; InterPro; IPRO03587; Hedgehog_FintN.

DR; InterPro; IPRO06141; Interin.

DR; InterPro; IPRO06142; INTEIN.

DR; InterPro; IPRO0442; Interin_endonuc.

DR; InterPro; IPRO04578; Pol2.

DR; Pfam; PF00136; DNA_Pol_B_1.

DR; Pfam; PF03104; DNA_Pol_B_exo; 1.

DR; PRINTS; PR00105; DNAPOLB.

DR; PR00379; INTEIN.

DR; SMART; SM00305; HintC; 1.

DR; SMART; SM00306; HintN; 1.

DR; SMART; SM00486; P0LBC; 1.

DR; TIGRFAMS; TIGR01443; interin_Cterm; 1.

DR; TIGRFAMS; TIGR01445; interin_Nterm; 1.

DR; TIGRFAMS; TIGR05592; Pol2_1.

DR; PROSITE; PS00116; DNA POLYMERASE_B; 1.

DR; PROSITE; PS05018; INTEIN_C_TER; 1.

DR; PROSITE; PS50819; INTEIN_ENDONUCLEASE; 1.

DR; PROSITE; PS50818; INTEIN_N_TER; 1.

KW; Transferase; DNA-directed DNA polymerase; DNA replication; DNA-binding; Autocatalytic cleavage; Protein splicing;

KW; Complete Proteome.

KW; CHAIN 1 492 DNA POLYMERASE, 1ST PART (POTENTIAL).

FT; CHAIN 493 952 PHO POLYMERASE, 2ND PART (POTENTIAL).

FT; CHAIN 953 1235 DNA POLYMERASE, 2ND PART (POTENTIAL).

FT; CHAIN 1235 AA; 143086 MW; 73CC7AA14873C84 CRC64;

FT; SEQUENCE 574: Conservative 107; Mismatches 89; Indels 464; Gaps 5;

Query Match 71.7%; Score 2886; DB 1; Length 1235;

Best Local Similarity 46.5%; Prod. No. 2e-149; Result 9

Matches 574; Conservation 107; Mismatches 89; Indels 464; Gaps 5;

Qy 1 MIFDDYTKDGPPIRKKENGFKIYDHPHQPIYALLKDDSDIDEIKAIKERHG 60

Qy 1 MILDADYTEDGPPIRKKENGFKIYDHPHQPIYALLRDSIDEIKKITAQRHG 60

Qy 61 KIVVVDAVKKKKFLGROEVWKLFLHFDQVPAKIRKIREPAVVDIYEDIPPAKRY 120

Qy 61 KIVVVETEKIQRFLGRIYEVWKLFLHFDQVPAKIRKIREPAVVDIYEDIPPAKRY 120

Db 121 LIDKGLIPMEGDEELKLMADFIEFYHEGDEFGKGETIMISTADEEEARVITWNIDLPY 180

Qy 121 LIDKGLIPMEGNEKLTFLAVDIELYHEGEEFKGPVIMISTADEEAKVITWNIDLPY 180

Db 181 VDVVSNERENIKRPFQIVREKDPVLYTNGDNFDLPLTIKRAEKGTLIGDRKHPE 240

Db 181 VEVVSEERENIKRFLRVIKEKDPDVITNGDNFDLPLKRAEKGTLIGKLLGRD--NSE 238

Qy 241 PKHRYMGDSPAVETIKRHFIDLPYTRTINLPYTYLEAVYAVLGTKTSKUGAEIAAI 300

Db 239 PKMQKNGDSLAVTRRIHDLFPVTRTINLPYTYLEAVYAVLGTKTSKUGAEIAAI 298

Qy 301 WETEESMKLQAQSMEDARATYELGEKEFFMEEAIAKLGQSYWVDSRSSTGNLVEWYL 360

Db 299 WETGSELERVKYMSMEDAKVYELGEKEFFMEEAQLRQGPWVDSRSSTGNLVEWYL 358

Qy 361 RVAYENELAPNKDEKEYERRLRLSTLGYWKEPERGLWENTYIDFRCLYSIIWTHN 420

Db 359 RKAYERNELAPNKDEKEYERRLRLSTLGYWKEPERGLWIVSDFRSLYPSIITIWHN 418

Qy 421 VSPDPLERESCKNDVAP1VYREFCDPFGFISIQLGELITIRQEKKMKATIDPIKK 480

Db 419 VSPDPLNREGCEEVYAPVKGHRICKDFFGFIPLLGOLLEERQKIKERMKESKDPIVEKK 478

Qy 481 MLDYRCARVL-----

Db 479 LLDYRORAIKILANSILPDEWLPIVENEKVRFYKIGDFIDREIBENAERVKRDGETEILE 538

Qy 492 -----

Db 539 VRDLKALSFNRETKSKESELKVYKALLRHYSGKYSIKLKGRRKITSGHSILPSVTKNGKL 598

Qy 492 -----

Db 599 VKYRGDELKPGDLYVVPGRKLPESKQVINVELLKLPIBETTSNIVMIPVKGRKNFK 658

Qy 492 -----

Db 659 GMLKLTLYWIFGEGERBRTAGYLKHLERLQYVVKURGEVLDNESSLKRYKLYTLLKN 718

Qy 492 -----

Db 719 LKYNGNSRAYMEFNSLRLDVYSLMPIEELKEMWIGEPRGPKIGTIDVDDSPAKLGLYYI 778

Qy 492 -----

Db 779 SSGDVEKDVKPHSKDQVNLIEDIAKLAELKFQVGRGKVRGRGYIEVSGSKHATFVLAEGKR 838

Qy 494 -----

Db 839 IPEPIFTSPMDIKVAPFLKGNGNAEELTPSTKSELNLINLNSIGVSDIKIHEKGV 898

Qy 494 -----

Db 899 YRVYINKKESSNDIVDSVSEIEVEKYEGLYDLSVEDENFLVGFLGLYAHNSTYGY 958

Qy 501 GYPRKARYSKCEAECSTAWGRHYIEMTIKEEKFGFKVLYADTDFGYATIPEGKP-ETI 559

Db 959 GYAKARWCKECAEBSVIAWGRQYIDVYRELAR-GRFKVLYIDTDFGYATIPEGVQWEEV 1017

Qy 560 KCKZKEFELKYINSKLPGLLEBYGFYLRGFFVAKKRYAVIDBEGRITTRGEVVERDWS 619

Db 1018 KZRALEFVDYDINSRKLPGTYELEYEGFARGFFVTKKYLIDBEGKIVTRGLEVVERDWS 1077

Qy 620 EIAKETQAKYLEALKEDSVEKAYEVYDVKVBEIAKQVPLFKEVTHBQITDSEYKAI 679

Db 1078 EIAKETQAKYLEALKHGNVEAKVKEVYDVKVTEKLTNTVEPPVLEIETQTRPNEYKAI 1137

Qy 680 GPHYAIKRAAKLKVPRGTTISYVILRGSKGKISDRVILLSEYDPKKHHKTYDPPDYYIENQ 739

Db 1138 GPHYAVAKLIMARSGKVKPGMVQYIVLRGDPGSKRAISIEFDPRKHKTDAEYYIENQ 1197

Qy 740 VLPATYRILLEAFGTRKEDLKYQSSKQVGILDWIK 773

Db 1198 VLPAVERILKAFGTYREDLWRQKTKQVGLGAWIK 1231

RESULT 9

DPOL THEFM STANDARD; PRT; 1523 AA.

ID P74918; AC P74918; DT 15-DEC-1998 (Rel. 3.7, Created)

DT 15-DEC-1998 (Rel. 3.7, Last sequence update)

DT 28-DEC-2003 (Rel. 4.1, Last annotation update)

DE DNA Polymerase (EC 2.7.7.7) (Pol Tfu) (Contains: Endonuclease PI-Tfu; DE (Tfu pol-2 (intein); Endonuclease PI-Tfu) (EC 3.1.-.-))

DE (Tfu pol-2 (intein); Endonuclease PI-Tfu) (EC 3.1.-.-))

GN PCL.

OS Thermococcus funicolans. Archaea; Buryarchaeota; Thermococci; Thermococcaceae;

OC Thermococcus; Thermococcus.

NCBI_TaxID=46540;	FT CHAIN 407	766	ENDONUCLEASE PI-PII.
RN [1]	FT CHAIN 767	900	DNA POLYMERASE, 2ND PART.
RN SEQUENCE FROM N. A.	FT CHAIN 901	1282	ENDONUCLEASE PI-PII.
RN STRAIN=ST557;	FT CHAIN 1283	1523	DNA POLYMERASE, 3RD PART.
RC Cabon M.; Querellou J.;	SQ SEQUENCE 1523 AA;	175917 MW;	7A2AC8236BF225F5 CRC64;
RA Submitted (FEB-1996) to the EMBL/GenBank/DDBJ databases.			
RT [2]			
RN CHARACTORIZATION OF INTEINS.			
RC STRAIN=ST557;	Query Match 69.8%; Score 2809; DB 1; Length 1523;		
RN MEDLINE=20112788; PubMed=10644683;	Best Local Similarity 39.0%; Pred. No. 4.1e-115;		
RX RT Saves 1.., Ozanne V., Dietrich J., Masson J.-M.;	Matches 593; Conservative 86; Mismatches 91; Indels 752; Gaps 4;		
RX RT "Inteins of the Thermococcus fumicolaans DNA polymerase are endonucleases with distinct enzymatic behaviors.",	Qy 1 MIFDTDXITKDGTGPPIRIFKKGKENGKPEKIELDPHFQPYTIALKKDDSAIDEIKAIKGGERHG 60		
RX RT Biochem. 275:2335-2341 (2000).	Db 1 MILDTDYTEDGPIVTFKKGKENGKPEKIELDPHFQPYTIALKKDDSAIDEIKAIKGGERHG 60		
CC -!- FUNCTION: PI-Tfull recognizes and cleaves a minimal sequence of 16 base pairs (bp) on supercoiled DNA with either Mn(2+) or Mg(2+) as cofactor. It cleaves linear DNA only with Mn(2+) and requires a 19 bp minimal recognition sequence. The optimal temperature for	Qy 61 KIVRVDVAKVKKKFLGRDVEWKKLIFEPHQDQVPAIGKIRHPAVDIFYDIPFAKRY 120		
CC activity is 70 degrees Celsius.	Db 61 TTIVRVRAGKVKKKFGRPIEWKKLYTHPQDQVPAIGKIRHPAVDIFYDIPFAKRY 120		
CC -!- FUNCTION: PI-Tfull is a highly active homing endonuclease using Mg(2+) as cofactor. Its minimal recognition and cleavage site is 21 bp long either on linear or circular DNA substrates. Its endonuclease activity is strongly inhibited by the 3' digestion product, which remains bound to the enzyme after the cleavage reaction. The optimal temperature for activity is 70 degrees Celsius.	Qy 121 LIDKGKLIPIMEGDBEKLKMAFDETFHGDGEGKGKELIMISYADEBEPARVTKWNLDPY 180		
CC -!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate + {DNA} (N).	Db 121 LIDKGKLIPIMEGDBEKLKMAFDETFHGDGEGKGKELIMISYADEBEPARVTKWNLDPY 180		
CC -!- PTM: THIS PROTEIN UNDERGOES A PROTEIN SELF SPLICING THAT INVOLVES A POST-TRANSLATIONAL EXCISION OF THE TWO INTERVENING REGION (INTEINS) FOLLOWED BY PEPTIDE LIGATION.	Qy 181 VDVVSNEBEMIKRFVQIVREKDPDVLITYNGDNFDYLPIKRAEKLGVTLIGRDEHPE 240		
CC -!- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-B FAMILY.	Db 181 VDVVSNEBEMIKRFVQIVREKDPDVLITYNGDNFDYLPIKRAEKLGVTLIGRDEHPE 238		
CC -!- SIMILARITY: IN THE INTEIN SECTION; BELONGS TO THE HOMING ENDONUCLEASE FAMILY.	Qy 241 PKTHRGDPSFAVBTIKGRTHFDLFPVYVRTINLPTYLEAVYEAUGKTKSKLGAEEAIAI 300		
CC	Db 239 PKTQRMGDRFAVEVKGRTHFDLFPVYVRTINLPTYLEAVYEAIFQPKVYAEIAQA 298		
CC	Qy 301 WETEBSMRKLAQSYMSMEDARATEYLGKFPMEALAKLIGOSWVWDYRSSTGNYLWYLL 360		
CC	Db 299 WETGEGLERRVATSMEDAKVTELGEGFPMEALQSLVQGFWDYSRSTGNYLWYLL 358		
CC	Qy 361 RVAAYERNELAPNPKPDEBEYRRRLRRTLYLGGYVKEPBRGLMENITYLDFRC----- 410		
CC	Db 359 RKAYERNELAPNPKPDEBEYRRRLRRTLYLGGYVKEPBRGLMENITYLDFRC----- 410		
CC	CC Between the Swiss Institute of Bioinformatics and the ENBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).	Qy 411 -----	410
CC	Db 478 TAIRDLSLAKSFLTKVYKGKLTITPLFEKIGKIEREDVPEEILKGELAGLIAEGLTLRK 537		
CC	Qy 411 -----	410	
CC	Db 538 DVEYFDSSRGKGRVSHQYRVEITVGAQEDFQRRIYIIFERLFGTVPFSYRKNTNTAIF 597		
CC	Qy 411 -----	410	
CC	Db 598 KVAKCEVYLVRVREIMDGIENLHAPSYLRGFPBGDSVNVKVRKTVVNNQGTNNEWKIEVVS 657		
CC	Qy 411 -----	410	
DR EMBL: Z69882; CAA9378.1; -.	Db 658 KLLNKG1P1PHRRTYTYDTERKTMTHILEIAGRDLILFQTTIVFISTERNMALEAR 717		
DR HSSP; P56639; 1TGO.	Qy 411 -----	420	
DR REBASE; 4400; PI-Tfull.	Db 718 NREVNRLLENNAAFTLADFTAKTEYYKGKVVYDLTBLGTPYFANGLTHNSLYPSIHN 777		
DR REBASE; 4501; PI-Tfull.	Qy 421 VSPDTLREGCKNVDYAPIVGYKFCCKDFGPGRPSIIGELITMRQETKKNKATPIEKK 480		
DR InterPro; IPR006172; DNA_Pol_B_dom.	Db 778 VSPDTLNRREGCCBEYDAPQSHRCFDGPGRPSIIGELDUDEROKVTKHAKATVDEIEKK 837		
DR InterPro; IPR006133; DNA_Pol_B_exo.	Qy 481 MLDYRQAVKVLHANSVYIGMYPKAWYSECAESVTAAGPHYIEMTIKEBKGFKVL 540		
DR InterPro; IPR003586; Hedehog_HintN.	DR PROSITE; PS50818; INTEIN_CTER; <u>2</u> ; FALSE_NEG.		
DR InterPro; IPR003587; Hedehog_HintN.	DR PROSITE; PS50819; INTEIN_CTER; <u>2</u> ; FALSE_NEG.		
DR InterPro; IPR006141; Intein.	DR PROSITE; PS50819; INTEIN_CTER; <u>2</u> ; FALSE_NEG.		
DR InterPro; IPR006142; Intein.	DR PROSITE; PS50819; INTEIN_CTER; <u>2</u> ; FALSE_NEG.		
DR InterPro; IPR004042; Intein_endonuc.	DR TIGRFAMS; TIGR01443; intein_Cterm; 2.		
DR Pfam; PF001336; DNA_Pol_B_2.	DR TIGRFAMS; TIGR01445; intein_Nterm; 2.		
DR Pfam; PF01104; DNA_Pol_B_exo; 1.	DR TIGRFAMS; TIGR01445; intein_Nterm; 2.		
DR PRINTS; PR00379; INTEIN.	DR PROSITE; PS50819; INTEIN_NTER; <u>2</u> .		
DR SMART; SM00305; HintC.	DR PROSITE; PS50816; DNA_POLYMERASE_B;		
DR SMART; SM00306; HintN; 2.	DR PROSITE; PS50818; INTEIN_CTER; <u>2</u> ; FALSE_NEG.		
DR SMART; SM00486; PolBc; 1.	DR TIGRFAMS; TIGR01443; intein_Cterm; 2.		
DR TIGRFAMS; TIGR01445; intein_Nterm; 2.	DR TIGRFAMS; TIGR01445; intein_Nterm; 2.		
DR TIGRFAMS; TIGR01445; intein_Nterm; 2.	DR PROSITE; PS50819; INTEIN_NTER; <u>2</u> .		
DR Transferase; PS50817; INTEIN_NTER; 2.	DR PROSITE; PS50817; INTEIN_NTER; 2.		
KW DNA-binding; Hydrolase; Endonuclease; DNA replication;	DB 838 LDYRQAVKVLHANSVYIGMYPKAWYSECAESVTAAGPHYIEMTIKEBKGFKVL 897		
KW Protein_splicing; Intron_mining; Magnesium; Manganese;	Qy 541 YAD-----	543	
KW DNA POLYMERASE, 1ST PART.	DB 898 YADSVTGDTEVTRRGRRIEFVPIKLFERVDRVQEKEYCVLGVWEALTLDNQRGLVWK 957		

RL	Nucleic Acids Res.	26:4409-4412 (1998).
CC	-1- FUNCTION: INTEIN ENCODED ENDONUCLEASES ARE THOUGHT TO MEDIATE	
CC	INTEIN MOBILITY BY SITE-SPECIFIC RECOMBINATION INITIATED BY	
CC	ENDONUCLEASE CLEAVAGE AT THE "HOMING SITE" IN GENE THAT LACK THE	
CC	INTEIN. PT-PROT RECOGNIZES 5'-GATTGATGCTGTGACC-3' AND PI-PROT	
CC	RECOGNIZES 5'-CAGCTACTACGGTAC-3'. BOTH ARE THERMOSTABLE.	
CC	+ [DNA] (N).	
CC	-1- PTM: THIS PROTEIN UNDERGOES A PROTEIN SELF SPLICING THAT INVOLVES	
CC	A POST-TRANSLATIONAL EXCISION OF THE INTERVENING REGION (INTEIN)	
CC	FOLLOWED BY PEPTIDE LIGATION (POTENTIAL).	
CC	-1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-B FAMILY.	
CC	-1- SIMILARITY: IN THE INTEIN SECTION, BELONGS TO THE HOMING	
CC	ENDONUCLEASE FAMILY.	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	
CC	the European Bioinformatics Institute. There are no restrictions on its	
CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement. (See http://www.isb-sib.ch/announce/	
CC	or send an email to license@isb-sib.ch).	
CC		
DR	EMBL; D29671; BAA06142; 2; -.	
DR	PDB; 1GCX; 2B-JAN-03.	
DR	REBASE; 3792; PI-PROT.	
DR	REBASE; 3793; PI-PROII.	
DR	InterPro; IPR000112; DNA_pol_B.	
DR	InterPro; IPR006134; DNA_pol_B_dom.	
DR	InterPro; IPR006133; DNA_pol_B_exo.	
DR	InterPro; IPR003586; Hedgenog_hintC.	
DR	InterPro; IPR003587; Hedgenog_hintCN.	
DR	InterPro; IPR006141; Interin.	
DR	InterPro; IPR006142; INTEIN.	
DR	InterPro; IPR004042; Intein_endonuc.	
DR	InterPro; IPR004578; PolB.	
DR	PF00136; DNA_Pol_B; 3; 1.	
DR	PFam; PF03104; DNA_Pol_B_exo; 1.	
DR	PRINTS; PR00379; INTEIN.	
DR	SMART; SM00305; HintC; 2.	
DR	SMART; SM00306; HintN; 2.	
DR	SMART; SM00486; POLB; 1.	
DR	TIGRFAMs; TIGR01443; interin_Nterm; 2.	
DR	TIGRFAMs; TIGR01445; interin_Nterm; 2.	
DR	TRIPOSITE; TIGR00592; PolB; 2.	
DR	TRIPOSITE; PS00116; DNA_Polymerase_B; 1.	
DR	TRIPOSITE; PS50818; INTEIN_C_TER; 2.	
DR	TRIPOSITE; PS50819; INTEIN-ENDONUCLEASE; 2.	
KW	Transferase; DNA-directed DNA polymerase; DNA replication;	
KW	Protein splicing; Hydrolase; Endonuclease; Autocatalytic cleavage;	
FT	CHAIN 1 406 DNA POLYMERASE, 1ST PART.	
FT	CHAIN 407 766 ENDONUCLEASE PI-PRO I.	
FT	CHAIN 767 851 DNA POLYMERASE, 2ND PART.	
FT	CHAIN 852 1388 ENDONUCLEASE PI-PRO II.	
FT	CHAIN 1389 1671 DNA POLYMERASE, 3RD PART.	
SQ	SEQUENCE 1671 AA; 193490 MW; 4A17F1C120BB455 CRC64;	
Query	Query Match 69.1%; Score 2782; DB 1; Length 1671;	
Query	Best Local Similarity 36.0%; Pred. No. 1.3e-143;	
Matches	Matches 602; Conservative 82; Mismatches 86; Indels 900; Gaps 4;	
Db	1 MIFDDPYITDGPKPIIRPKCENGEEFKIQLGRDVEWKLIIFBPQYIYALKDOSAIDEKAIKGERRG 60	
Db	1 MILDPTYITDGPKPIIRPKCENGEEFKIQLGRDVEWKLIIFBPQYIYALKDOSAIDEKAIKGERRG 60	
Qy	61 KIVVVDAVKKKKFGLGRDVEWKLIIFBPQYIYALKDOSAIDEKAIKGERRG 120	
Db	61 TVVTVKRVVERVQKFLGRPEVYKLYFTHPQDVA-TRDKIREHPAVIDYDIFFAKRY 120	
Qy	121 LIDKGLIIPMEGDEBKLMADISYHGGDEGKGETIMISYADEBEARVITWKIDLPY 180	

Db	121	LIDKGLVPMEGDBELKMLAFDIEITLYHEGEEPAEGPLIMISVADEBCARVITWKNDLPLV	180	494	Qy
Qy	181	VDYVSNEREMIKRFVOLIVTREKOPDVLYTNGDNFDLPLIKRAEKUGVTLIGRKEHPE	240	493	Db
Db	181	VDYVSNEREMIKRFVOLIVTREKOPDVLYTNGDNFDLPLIKRAEKUGVTLIGRKEHPE - E	238	493	Qy
Qy	241	PKIHRMGDSFAVYBTKGRHEDLFPVYRTRINLPTYLBAVYBALKTKSGLGAEIAAI	300	494	Db
Db	239	PKIQRMGDRFAVYBVKGRHFDLFPVYRTRINLPTYLBAVYBALKTKSGLGAEIAAI	298	494	Qy
Qy	301	WETEFSMKKLQYSMDSARATYLGKEFPMMAELAKLIGQSWMDRSRSGNLYEWLL	360	494	Db
Db	299	WETEFSMKKLQYSMDSARATYLGKEFPMMAELAKLIGQSWMDRSRSGNLYEWLL	358	494	Qy
Qy	361	RVAYERNLAPKPDDEFYRRRLRTTLLGGTYKEPERGLWENITYLDFC	410	544	Db
Db	359	RVAYERNLAPKPDDEFYRRRLRTTLLGGTYKEPERGLWENITYLDFC	417	1438	Qy
Qy	411	411	410	1378	Db
Db	418	GKGIINISSEVQEGDYVLGIDGKQRVTKWEDYKGELVNINGLKCTPNPKLPUVTKNBRQ	477	1498	Qy
Qy	411	411	410	1497	Db
Db	478	TRIRDLSLAKSFLTKVKQKIIITPLFYIGRATSENIPEEVTLKGELAGILLAEGTLRK	537	664	Qy
Qy	411	411	410	604	Db
Db	538	DVEYFDSSRKRKREISHQYRVEITIGDDEEFDRITYIFERLFGITPSISERKGNTAVTL	597	1558	Qy
Qy	411	411	410	724	Db
Db	598	KVAKKNTYLVKVBIMDNNEISLHAPSVLRGFFGDGSYNRVRRSIVATQGTNNEWKIKLVS	657	1618	Qy
Qy	411	411	410	1618	Db
Db	658	KILSQLGIIPHOTTYQYQENGKDRSRYLEITGKDGLLFTQNLIGTSEKRNALLNTKAS	717	16-OCT-2001	QDOL THEG8
Qy	411	411	410	16-OCT-2001	AC Q9H864;
Db	718	QREMNNTLENGFYRLSEENVSTYEYKGKVYDITLEGTPYYFANGILTHNSLYPSIITEN	777	16-OCT-2001	DT 16-0CT-2001 (Rel. 40, Created)
Qy	421	VSPDTLLEBGKRYDVAIYGKXPCDFPGFIPSILQBLITMREQITKMKATDIERK	480	28-FEB-2003	DT 28-02-2003 (Rel. 41, Last sequence update)
Db	778	VSPDTLNEGCFKEYDVAQVGHFRCKDFPGFIPSILQBLITMREQITKMKATDIERK	837	DB DNA Polymerase (EC 2.7.7.7) [Contains: Endonuclease PI-TspG8I	
Qy	481	MLDYRORAVKLA-----	493	DE (EC 3.1.1.-) (TSP-GE8 pol-1 intein); Endonuclease PI-TspG8I	
Db	838	LLDYRQAKILANSILPEEWLPLVSEHVFVIGELIDRMMEENAGKVREGETEVLE	897	DN POL OR POL-1.	
Qy	494	897	GN POL OR POL-1.		
Db	898	VSGLEVPSFNRRNTKAELKRVKALIRHDYSGKVYTLRKSRRKITSGHLSFSVRNGEL	957	OS Thermococcus sp. (strain GBE).	
Qy	494	957	OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococaceae;		
Db	958	VEVTGDELKPGDIAVAPRPLERPERNHVNLVYELLIGTPEEETLIVMTIPVKGRKKNFFK	1017	OC Thermococcus	
Qy	494	1017	NCB_105583; RN [1]		
Qy	1018	GMRLTRWTFGEERPRARRYLRLHELGYVRLKXIGEYVLDWDSLKNTVRLYALVEN	1077	RP SEQUENCE FROM N. A.	
Db	494	1077	RA Berbier J.J.E., Cambon M.A., Lescoueur F., Barbier G.;		
Qy	1078	VRYNGKREYLVBNFSTDAGIMPLKELKEMKIGTLNGFMRKLIEVEDSLAKLGLGYV	1137	RT "Thermococcales taxonomy and phylogeny based on the comparative use of	
Db	494	1137	RT 16S rDNA, 16S-23S rDNA intergenic spacer and Family B DNA polymerase		
Qy	1078	493	RT genes";		
Db	494	493	RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.		
Qy	494	493	CC -!- FUNCTION: IN ADDITION TO POLYMERASE ACTIVITY, THIS DNA POLYMERASE		
Db	494	493	CC EXHIBITS 3-, TO 5-, EXONUCLEASE ACTIVITY (BY SIMILARITY).		
Qy	494	493	CC -!- FUNCTION: PI-TSPG8I AND PI-TSPG8II ARE ENDONUCLEASES		
Db	958	VEVTGDELKPGDIAVAPRPLERPERNHVNLVYELLIGTPEEETLIVMTIPVKGRKKNFFK	1017	CC -!- (POTENTIAL) ACTIVITY: N deoxyribonucleoside triphosphate = N diphosphate	
Qy	494	1017	CC -!- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-B FAMILY.		
Db	1078	VRYNGKREYLVBNFSTDAGIMPLKELKEMKIGTLNGFMRKLIEVEDSLAKLGLGYV	1137	CC -!- SIMILARITY: IN THE INTIN SECTION; BELONGS TO THE HOMING	
Qy	494	1137	CC -!- ENDONUCLEASE FAMILY.		
Db	1138	SEGYARKQRNPKNQWSYKLYNEDPEVDDMERLASRFFGKVRGRNYYEIPKKGYLL	1197	CC This SWISS-PROT entry is copyright. It is produced through a collaboration	
Qy	494	1197	CC between the Swiss Institute of Bioinformatics and the EMBL outstation -		
Db	1198	FENMGVLAENKRIPEFTVFTPKGVRLAFLEGYFIGDDVHPNKRDLRLTRSELLANQLV	1257	CC the European Bioinformatics Institute. There are no restrictions on its	
Qy	494	1257	CC use by non-profit institutions as long as its content is in no way		
Db	1198	1257	CC modified and this statement is not removed. Usage by and for commercial		
Qy	494	1257	CC entities requires a license agreement (See http://www.isb-sib.ch/announce/		
Db	1198	1257	CC or send an email to licensing@isb-sib.ch).		

DR	AJ250333; CAC12850.1; -.	Qy	494
DR	IP6689; 1TGO.	Db	538
DR	InterPro; IP006172; DNA_Pol_B.	Qy	494
DR	InterPro; IP006134; DNA_Pol_B_dom.	Db	597
DR	IP66123; DNA_Pol_B_exo.	Qy	493
DR	InterPro; IP003516; Hedgehog_hintC.	Db	598
DR	InterPro; IP003558; Hedgehog_hintN.	Qy	493
DR	IP66141; Intein.	Db	657
DR	InterPro; IP004012; Intein_endonuc.	Qy	493
DR	IP004026; PolB.	Db	658
DR	InterPro; IP004578; Pol2.	Qy	493
DR	Pfam; PF00136; DNA_Pol_B_3.	Db	717
DR	Pfam; PF0104; DNA_Pol_B_exo; 1.	Qy	493
DR	PRINTS; PR0379; INTEIN.	Db	718
DR	SMART; SM00305; HintC.	Qy	493
DR	SMART; SM00306; HintN.	Db	777
DR	SM00486; PolBc.	Qy	494
DR	TIGRFAMS; TIGR0143; interin_Cterm; 2.	Db	778
DR	TIGRFAMS; TIGR0145; interin_Nterm; 2.	Qy	494
DR	TIGRFAMS; TIGR0522; Pol2; 2.	Db	838
DR	PROSITE; PS00116; DNA_POLYMERASE_B; FALSE_NEG.	Qy	493
DR	PROSITE; PS00818; INTEIN_CTER; 2.	Db	898
DR	PROSITE; PS00819; INTEIN_ENDONUCLEASE; 2.	Qy	494
DR	Transferease; DNA-directed DNA polymerase; DNA replication;	Db	944
KW	KW	Qy	493
KW	Protein_splicing; Hydrolyase; Nuclease; Exonuclease; Multifunctional enzyme;	Db	944
KW	DNA-binding; DNA-directed DNA polymerase; Endonuclease; Intron homing.	Qy	493
FT	1 PROSITE; PS00116; DNA_POLYMERASE; 3.	Db	944
FT	CHAIN 491 ENDONUCLEASE PI-TSPG8II.	Qy	494
FT	CHAIN 492 1026 DNA POLYMERASE PI-TSPG8II.	Db	944
FT	CHAIN 1027 1075 ENDONUCLEASE PI-TSPG8II.	Qy	494
FT	CHAIN 1076 1464 ENDONUCLEASE PI-TSPG8II.	Db	958
FT	CHAIN 1465 1699 DNA POLYMERASE; 3RD PART.	Qy	494
FT	SEQUENCE 1699 AA; F330B4351F0B12D3 CRC64;	Db	1018
Qy	Query Match 68.8%; Score 2769.5; DB 1; Length 1699;	Qy	494
Qy	Best Local Similarity 35.5%; Prod. No. 6.5e-143; Matches 83; Mismatches 83; Indels 927; Gaps 4;	Db	1078
Db	1 MIFDTDYIKDQGPIIRFKKENGEFFKLELDPHQFQPTVYALLQDDAIDETKAIGKRGH 60	Qy	544
Db	1 MIFDTDYIKDQGPIIRFKKENGEFFKLELDPHQFQPTVYALLQDDAIDETKAIGKRGH 60	Db	1078
Qy	61 KIVRVVDAVVKKKFLGRVEWKLIFPQDPAFLGKIREHPAYTDIYEVDIFAKCY 120	Qy	544
Db	61 TVVVKYRAVKVKKFLGRPIEWKLIFPQDPAFLGKIREHPAYTDIYEVDIFAKCY 120	Db	1138
Qy	121 LIDKGLIPMEGDEBKLMAPDIFTYHEGDEFGKGEIIMISADEEFARYTWNIDLPY 180	Qy	544
Db	121 LIDKGLIPMEGDEBKLMAPDIFTYHEGDEFGKGEIIMISADEEFARYTWNIDLPY 180	Db	1198
Qy	181 VDVVSNEREMIKRPFQVIREKDPVLYTNGNFDLPLIKRAEKGVTLLGRDEKHP 240	Qy	544
Db	181 VDVVSNEREMIKRPFQVIREKDPVLYTNGNFDLPLIKRAEKGVTLLGRDEKHP 240	Db	1258
Qy	181 VDVVSNEREMIKRPFQVIREKDPVLYTNGNDFAYLKRSEKLGKVFILGRGS--E 238	Qy	544
Qy	241 PKIHRMGDSFAVEIKGRTHFDLPPVVRTRNLPYTYEAVLICKTSKLGAEIIRAI 300	Qy	544
Db	239 PKIORMGDRPAVEYKRIIFDLPVTRINPPTYEAIEFGPKKEVYABIAITA 298	Qy	544
Qy	301 WETEWSMKGKLAQSMEDARATYELGKEFFPMFAELAKLIGOSWVDSRSSTGNIVEWYIL 360	Db	1378
Db	299 WETEWSMKGKLAQSMEDARATYELGKEFFPMFAELAKLIGOSWVDSRSSTGNIVEWYIL 360	Qy	544
Qy	361 RVAYERNELAQNPKDPEEYRRRLRTTYLGGYTYKEPERGLWENITYDLFCLYPSITWHN 420	Db	1318
Db	359 RKAYERNELAQNPKDPEEYRRRLRTTYLGGYTYKEPERGLWENITYDLFCLYPSITWHN 417	Qy	544
Qy	421 VSPDPTLREBCKNQYDVAAP-IVGYRCKDQPGFPTISIGELITMREKKMKATIDPEKK 480	Db	1498
Db	418 VSPDPTNREGKEYDVAPOVGHKEFCKDQPGFPTISIGELITMREKKMKATIDPEKK 477	Qy	637
Qy	481 MLDYQRAYKLHA--	Db	1553
Db	478 LLDRYQRRAIKILANSILPDEWLPILVNGRLKLYRIGFVDNTMKKGQPLENDGTEVLEVS 537	Qy	1617

Qy 697 REGTISYIVLRSKISDRVILLSEYDPKHKYDPPDYYTENQVLPFAVRLLEAFGYRKE 756
 Db 1618 RPTVISYIVLKGSGRIGDRAFPDFDPKHKYDABYYIENQVLPFAVRLLEAFGYRKE 1677

Qy 757 DRYQSSKQVGLDAWIK 773
 Db 1678 DRYQTKQVGLGAWIK 1694

RESULT 12

ID DPOL THIHY STANDARD; PRT; 1668 AA.

AC Q9HH05 [1] 40. Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE DNA Polymerase (EC 3.1.17.7) [Contains: Endonuclease PI-ThiI]. (EC 3.1.1.-.)

DE (BC 3.1.1.-.) (ThiI Pol-1 intein); Endonuclease PI-ThiI (EC 3.1.1.-.)

DE (ThiI Pol-2 intein) (Fragment).

GN POL.

OS Thermococcus hydrothermalis.

OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;

OC Thermococcus

OX NCBI_TaxID=46539;

RN [1]

RP SBOQUENCE FROM N. A. Barbier G.; Querellou J.J.E.; Cambon M.A.; Lesongeur P.O.; Barbier G.; RT Thermococcales taxonomy and phylogeny based on the comparative use of 16S rRNA, 16S-23S rDNA intergenic spacer and family B DNA Polymerase genes.", Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.

[2]

RP CHARACTERIZATION OF PI-THIY; MEDLINE=2012590; PubMed=11058140; RA Saves J., Bleaume H., Dietrich J., Masson J.-M.; "The Thi pol-2 intein of Thermococcus hydrothermalis is an isozyme of PI-ThiI and PI-ThiII endonucleases.", RT Isozyme of PI-ThiI and PI-ThiII endonucleases.", Nucleic Acids Res., 28:4391-4396 (2000).

CC -1- FUNCTION: IN ADDITION TO POLYMERASE ACTIVITY, THIS DNA POLYMERASE EXHIBITS 3', TO 5', EXONUCLEASE ACTIVITY (BY SIMILARITY).

CC -1- FUNCTION: PI-THIY AND PI-THIII ARE ENDONUCLEASES. PI-THIY CLEAVES THE INTEINLESS SEQUENCE OF THE THI DNA POL. GENE. IT REQUIRES A 21-BP MINIMAL RECOGNITION SEQUENCE.

CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate + (DNA) (N).

CC -1- PTM: THIS PROTEIN UNDERGOES A PROTEIN SELF SPLICING THAT INVOLVES A POST-TRANSLATIONAL EXCISION OF THE INTERVENTING REGION (INTEIN) FOLLOWED BY PROTEIN LIGATION (POETRY).

CC -1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-B FAMILY.

CC -1- SIMILARITY: IN THE INTEIN SECTION, BELONGS TO THE HOMING ENDONUCLEASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; AJ245819; CAC18555.1; -.

DR HSSP; P56683; ITC0.

DR REBASE; 4832; PI_ThiI.

DR InterPro; IPR006172; DNA_Pol_B.

DR InterPro; IPR006134; DNA_Pol_B_dom.

DR InterPro; IPR006133; DNA_Pol_B_exo.

DR InterPro; IPR003586; Hedgehog_HintC.

DR InterPro; IPR003587; Hedgehog_hintN.

DR InterPro; IPR006141; Intein.

DR InterPro; IPR006142; Intein.

DR InterPro; IPR004042; Intein_endonuc.

DR InterPro; IPR004578; Pol2.

DR Pfam; PF00136; DNA_Pol_B; 3.

DR Pfam; PF03104; DNA_Pol_B_exo; 1.

DR PRINTS; PR00379; INTEIN.

DR SMART; SM00305; HintC; 2.

DR SMART; SM00306; HintN; 2.

DR TIGRFAMS; TIGR0143; Intein_Cterm; 2.

DR TIGRFAMS; TIGR0144; Intein_Nterm; 2.

DR TIGRFAMS; TIGR000522; Pol2; 2.

DR PROSITE; PS00116; DNA_POLIMERASE_B; FALSE_NEG.

DR PROSITE; PS55818; INTEIN_C_TER; 2.

DR PROSITE; PS55819; INTEIN_N_TER; 2.

KW Transferase; DNA-directed DNA Polymerase; DNA replication;

KW DNA-binding; Hydrolase; Nuclease; Exonuclease; Multifunctional enzyme;

KW Protein splicing; Autocatalytic cleavage; Endonuclease; Intron homing.

PT NON_TER [1] 458 DNA POLYMERASE, 1ST PART.

FT CHAIN [1] 458 ENDONUCLEASE PI-THIY (POTENTIAL).

FT CHAIN [1] 995 DNA POLYMERASE, 2ND PART.

FT CHAIN [1] 1044 ENDONUCLEASE PI-THIY (POTENTIAL).

FT CHAIN [1] 1433 DNA POLYMERASE, 3RD PART.

FT CHAIN [1] 1668 DNA POLYMERASE, 3RD PART.

SEQUENCE 1668 AA; 193319 MW; 5EBB805FFEDACRC64;

Query Match Score 63.0%; Score 2535.5; DB 1; Length 1668;

Best Local Similarity 33.6%; Pred. No. 3..3e-30;

Matches 560; Conservative 85; Mismatches 92; Indels 929; Gaps 4;

Qy 34 FQPYIALLKDDAIDEIKAKGGERHKGTVVYDAVYKVKCFGLGRDVEVWLLFIFEFQDV 93

Db 1 FEPYIALLKKDAAIEYVKKITAGRVRVVKYKRAEYVKKPLGRPEVWLLYTFHQDV 60

Qy 94 PALRKIREHPAVDIDYDIFPKARYLIDKGLIPMEGDEELKLNADFIETYHEGDEFG 153

Db 61 PAIRDEIRHSAVVDIYDIFPKARYLIDKGLIPMEGDEELKLNMSDIFTLYHEGEEFG 120

Qy 154 KGETIMIISYADEEVARVITWKNDLPIYDPIVSNEREMIKRFEVIVPEKDPVLTGNDN 213

Db 121 TGPMLMISADEGBARVITWKCDLPIYDPEVWSTEKNIKRFIKVKEVDPVLTGNDN 180

Qy 214 FDLPYLIKRAEKGKVGLTLLGRDKEHPKPIHMRGDSFAVEIKGRHIFDLPVYVRRTINLP 273

Db 181 FDFAYLKERKCEKIGIKFTFLRQDS-EPKIQMGRDFAVEYKGRIFDLPVYVRRTINLP 238

Qy 274 TTYLAEVAYEAVLGTKTSKLGAEFIAATWETEESMKLQIAQYSMEDATAYELKEGKFPPMEA 333

Db 239 TTYLAEVAYEAVFGTPKVYPEEITAWETGEGLBRVARYSMEDATKTYELGRBFPPMEA 298

Qy 334 ELAKLIGOSWDYSSRSGNLYEWYLAVTERNEALAPKDEEEFRRLTTLTYLGYVK 393

Db 299 QLSRLIGOSLWDYSSRSGNLYEWYLKAVERNEALAPKDERELARR-RGGYAGGYVK 357

Qy 394 EPERGLWNNITYLDFCLYPSIIVTWNVSPDTLERSCKSKNYDVAPIVGYKECKDFGFIP 453

Db 358 EPERGLWNNVIVDMLWPSIITITVSPDTNPSDFTNRGCKYDTAPVQHFKDFDSFIP 417

Qy 454 STIGELITMROBTKKKKATDPIEKMLDYQRQAKLHA----- 493

Db 418 SLIGALDDEQKIKRMAKIDPLKCLLDTQRAKIKLANSLLPEPWIPLYENGKVRLH 477

Qy 494 ----- 493

Db 478 RIGEFVDKLMETDSELYVRNGDTEVLEVRGIRALSDRKSKCARYVMVKAVIRHRYSGDV 537

Qy 494 ----- 493

Db 538 YEIVLGSGRRTYTEGHSLFAYGDFELREVTCGKIDGDLIAVPRVNLPECKERLNVL 597

Qy 494 ----- 493

Db 598 LLRRLPPEETGDLILTIPVKGRKNFVGMLRTLWISGEEKPRTRARRYLEHLEGIGYVR 657

Qy 494 ----- 493

370 APNKPDEEYRRRLRTYLGYYKEPPERGLMENITYLDRCLYPSITIHRVSPDTL--E 427
 413 SPQRPSSSYREPAKYYEGGYVREPKQIDIVSDFMSLPSLISNISPTVIE 472

Qy 428 RESCKNDVAPIGYKCFKDFGIFPILGELITIQRKQIKMATAKIDPIEK-----K 480
 Db 473 EKRENML-----GLIPTNLNEILSRKHKRNLK--DK1QKNEEDEEYS 516

Qy 481 MLDYRQRARVQLHANSYYGMPKARNYSKCEASVTAWRHYIEMTKEEKFQFVYL 540
 Db 517 RLEHEQSKIVLANSHYGGLAPMARYSDKCAEMYTGRLGKYIQTIEKAE-FGFKV 575

Qy 541 YADTDGFATIPECEKPTIKK-----AKEFLAYIINSKLPGLL 578
 Db 576 YADTDGFIAKWDYDQKLGKKEBENDSKDLSNLPKUSKEEBELITTKFLGINEELPGEM 635

Qy 579 ELEYEGPFLRGFFEYAKERAYAVIDEGRITTRGLEYVRDMSBIAKETAQKYLEALKEDS 638
 Db 636 ELEFEGHHTFRGLFVTRCKYALLEDDGHIVVKGLEYVRDWSNIAKDTQOAVERALLEDGD 695

Qy 639 VKEAVEVTKDVBEEIAKQVPLEKLYTHEQFTKDLSEYKALGPHVIAARRLAANGIKVRP 698

Db 696 VNLAKK1K1KNTIDNLKKGK1DNL1HTQ1LTKNTEBEYKSSTAPHEVAKKIKORDSVER 755

Qy 699 GPTISYIVLRGGSKISDPRVILSEYDPRKHKYDPPYI1ENQVLPVRLDEAFGTYRKEDL 758

Db 756 GPTISYIVKGGRSISRAELL-EY---AGDIDNQVLPVPRIMTSLGTSDEL 811

Qy 759 KYOSSKQVGLDAWL 772

Db 812 K-NSGKQFLQDFM 824

RESULT 14
 DEOL_ARCFU STANDARD PRT; 781 AA.
 AC 029753_1 ID DEOL_ARCFU STANDARD PRT; 781 AA.
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE DNA Polymerase (EC 2.7.7.7).
 GN POLB OR AF04947.
 OS Archaeoglobus fulgidus.
 OC Archaeoglobi; Archaeoglobales;
 OC Archaeoglobaceae; Archaeoglobus.
 OX NCBI TaxID:2234;
 RN [1] SEQUENCE FROM N. A.
 RC STRAIN:VC-16 / DSM 4304 / ATCC 49556;
 RRK M88941; Pubmed:3389475;
 RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.B.,
 Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
 Richard D.L., Kerlavage A.R., Graham D.E., Kyrides N.C.,
 Fleischmann R.D., O'Neil J., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
 Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Lottus B.,
 Reich C., McNeil L.K., Badger J.H., Girod A., Zhou L.,
 Overbeck R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
 Cotton M.D., Spriggs T., Artiach P., Kalne B.P., Sykes S.M.,
 Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
 Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
 Venter J.C.,
 RT "The complete genome sequence of the hyperthermophilic, sulphate-
 reducing archaeon Archaeoglobus fulgidus.";
 RL Nature 390:364-370 (1997).
 CC + (DNA) (N). Catalytic activity: N deoxyribonucleoside triphosphate = N diphosphate
 CC -! SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-B FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the FMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 DR EMB2; AB001070; AAB90741.1; -.
 DR HSS; PE63312; A69312.
 DR TIG; AB0497; -.
 DR InterPro; IPR006172; DNA_Pol_B.
 DR InterPro; IPR006134; DNA_Pol_B.
 DR InterPro; IPR006133; DNA_Pol_B.
 DR InterPro; IPR004578; Pol2.
 DR Pfam; PF00136; DNA_Pol_B.
 DR Pfam; PF03104; DNA_Pol_B.
 DR SMART; SM00486; DNAPOLB.
 DR TIGRFAMS; TIGR00932; Pol2; 1.
 DR PROSITE; PS00116; DNA_POLYMERASE_B; 1.
 KW Transferase; DNA-directed DNA Polymerase; DNA replication;
 KW DNA-binding; Complete proteome; DNA binding; Complete proteome;
 SQ SEQUENCE 781 AA; 8985 MW; 761C7B18FCC61B8B CRC64;
 Query Match 32.3%; Score 1301; DB 1; Length 781;
 Best Local Similarity 37.4%; Pred. No. 1.5e-63;
 Matches 305; Conservative 155; Missmatches 26; Indels 90; Gaps 21;
 Qy 2 IFDFTYRKGKPIRIRKENGKFIELDPHFQPYIALLYKDDSAIDEIKRIGGRHKG 61
 Db 8 LIDADYETIGKAVVR1WCKDQGIFTAYDNNFDPFYVIGYDE--DILKNAATSTRE 64
 Qy 62 IVRVVDAYKVKKKPLGRDVEVKLIFHPQDVPLSKIREHPAVIDYEDIPFAKRYL 121
 Db 65 VVILKSPFKAOLKLTGREVEG3IVYAHPOHYPKLRYDLSQFG--DVERADIPFAKRYL 121
 Qy 122 IDGL-----IPMEGD-----EIKLMAFIETTYHEG DEFG 153
 Db 122 IDDLACMDGIAEGERQGGVIRSYKIEKVERIPRMFPEKPMVFDCEMLSSPGMPEPE 181
 Qy 154 KGETIMISYADEBEARVITIWNIDLPYDVPVNSNREMIKRFPQIVREKDPPDVLITYNGDN 213
 Db 182 KDTTIVISVKINDDEII-----LTSDEERK1TSDFVLLKSYDPLIVGQNDA 230
 Qy 214 FDLPLYLTRAEKLGVTUILLGRDKEHPEPKHFMGDDPAVEFKGRHFDLFPVVRRTINLP 273
 Db 231 FDPDPYLRAERVNIPDVGROSNN--VVRGG--RPKTGRNLVNDLYDAMRSVDIK 284
 Qy 274 TYLPLAYEAVLIGKTA-SKLGAEERIAIAWETEESMVKLAQYSMEDDABATYBLGKFPPE 332
 Db 285 IKCLEVNAEFLGKTKIAEADIBARDIYRWSREK-BKVLNTYARODINTYLIAKEELPMPH 343
 Qy 333 AELAKLIGQSVDYDSRSSTSGNLYVETWYLLRAXERNEELAPNKPDEEYRRRLRTTYGGV 392
 Db 344 YEUSKMRPLPVDTVRMGRKQYDWTWLLSEARKIGTIANPPEAH-----SYGAFY 396
 RA 393 KEPERGLMENITYLDFCLYPSIILVTHNVSPTDLEREGKNN--YDAPATVGYKFCDFPG 450
 RA 397 LEPERGLHENVACLDASMSYPSIMAFNISPDY--GCRDCYE-APEYGHKFKXSPDG 452
 RA 451 FBSILGFLITNRQEIKKRMK-ATIDPTEKMDYRORAVKLHANSYYGNGYPKARWYS 509
 RA 510 KECAESVTTAWGRHYIENTIKEEKFVKLYADTDFYATIPIGERPETYKAKFLKY 569
 RA 513 HPCAEATTAAGHAFIRTSAK-TAESNGFKVLYGDTOSITFVKAGMKTEDVDR---LIDK 567
 Qy 570 INSKLPGILEBYEGFLRGFFPVAKKRYAVIDEGRITTRGLEYVRDWSPIAKETOAKV 629
 Db 568 LHELP--IQLVDEYSAIIEVEKRYAGLTDGLLVKGLEVRGDWCLAKKQREV 625
 Qy 630 LPAIKEDDSVKEAVETVYDVEEIAKQVPELKVTHEQTKDLSVSKYAKGPHVAKRKL 689
 Db 626 IEVILKEKNPERALSIVDYLRIKESKVSLEVVYTKGLITKKEPSYEMQAHVAKLA 685

RESULT 15						
	DPOL_METHA	STANDARD;	PRT;	1634 AA.		
Qy	ID DPOL_METHA AC P0395;					-HKEYDPXX 735
	DT 01-NOV-1997 (Rel. 35, Created)					
Db	DT 15-DEC-1998 (Rel. 37, Last sequence update)					
	DT 28-FEB-2001 (Rel. 41, Last annotation update)					
Qy	DE DNA polymerase (EC 2.7.7.7) [Contains: Mja pol-1 intein; Mja pol-2 intein].					
Db	DE [1] RN POL_QR_M00885.					
	OS Methanococcus jannaschii.					
	OC Archaea: Euryarchaeota; Methanococci; Methanococcales;					
	OC Methanococciococaceae; Methanococcoides.					
	NCBI_TaxID=2195;					
	OX [1]					
	RP SEQUENCE FROM N.A.					
	RC STRAIN_JAL-1 / DSM 2661 / ATCC 43067;					
	RX MEDLINE=96337999; PubMed=8681087;					
	RA Built C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,					
	RA Sutton G.G., Blake J.A., Pitkäraid L.M., Clayton R.A., Gocayne J.D.,					
	RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,					
	RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,					
	RA Scott J.L., Geoghegan N.S.M., Weidman J.L., Fuerhmann J.L., Nguyen D.,					
	RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,					
	RA Cotton M.D., Rogers K.M., Hurst M.A., Borodovsky M.,					
	RA Klenk H.-P., Fraser C.M., Smith H.O., Kaine B.P., Venter J.C.,					
	RT "Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii." Science 273:1058-1073 (1996).					
	CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate					
	CC + {DNA} (N).					
	CC -1- PTM: PROTEIN UNDERGOES A PROTEIN SELF SPLICING THAT INVOLVES					
	CC A POST-TRANSLATIONAL EXCISION OF THE INTERVENING REGION (INTEIN)					
	CC FOLLOWED BY PEPTIDE LIGATION (POTENTIAL).					
	CC -1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-B FAMILY.					
	CC					
	CC This SWISS-PROT entry is copyright. It is produced through collaboration					
	CC between the Swiss Institute of Bioinformatics and the EMBL outstation					
	CC the European Bioinformatics Institute. There are no restrictions on its					
	CC use by non-profit institutions as long as its content is in no way					
	CC modified and this statement is not removed. Usage by and for commercial					
	CC entities requires a license agreement (See http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).					
	CC					
	DR U67532; AAB98899.1; -.					
	DR HSSP; P56683; 1TGO.					
	DR TIGR; MJ0885; -.					
	DR InterPro; IPR006172; DNA_Pol_B.					
	DR InterPro; IPR006134; DNA_Pol_B_dom.					
	DR InterPro; IPR006133; DNA_Pol_B_exo.					
	DR InterPro; IPR005186; Hedgehog_FaintC.					
	DR InterPro; IPR003587; Hedgehog_hintC.					
	DR InterPro; IPR006141; Intein.					
	DR InterPro; IPR006142; Intein_endonuc.					
	DR InterPro; IPR004578; Pol2_.					
	DR Pfam; PF00116; DNA_Pol_B_3.					
	DR Pfam; PF03144; DNA_Pol_B_exo; 1.					
	DR InterPro; PRO0379; INTEIN.					
	DR SMART; SM000105; HintC_2.					
	DR SMART; SM000106; HintN_2.					
	DR SMART; SM000107; HintN_3.					
	DR TIGRTRANS; TIGR01443; inrein_Cterm; 2.					